Controlling Customer Master Data Quality: Findings from a Case Study

Ehsan Baghi  
University of St. Gallen, ehsan.baghi@unisg.ch

Boris Otto  
University of St. Gallen, boris.otto@tu-dortmund.de

Hubert Oesterle  
University of St. Gallen, hubert.oesterle@unisg.ch

Follow this and additional works at: http://aisel.aisnet.org/confirm2013
Controlling Customer Master Data Quality: Findings from a Case Study

Ehsan Baghi
University of St. Gallen
ehshan.baghi@unisg.ch

Boris Otto
University of St. Gallen
boris.otto@unisg.ch

Hubert Oesterle
University of St. Gallen
hubert.oesterle@unisg.ch

Abstract
Data quality management plays a critical role in all kinds of organizations. High-quality data is one of the most important prerequisites for making strategic business decisions and executing business processes. In order to be able to assess data quality, ensure efficient process execution, and verify the effectiveness of data quality initiatives, data quality has to be monitored and controlled. This can be achieved by implementing a comprehensive controlling system for data quality. However, only few organizations have managed to implement such a system. This paper presents a single-case study describing the process of implementing a comprehensive data quality controlling system. The study focuses on controlling activities defined in the field of business management.

Keywords
Data Quality Management, Data Quality Controlling, Master Data Management, Master Data Quality
Controlling Master Data Quality: Findings from a Case Study

1. Introduction
One of the most important drivers of ERP software implementation in the last decade has been the need for a better quality of data and information in enterprises (Berndroder and Hampel, 2005). Data quality management (DQM), as a part of enterprise information management, plays a critical role in all kinds of organizations. Poor data quality can have a decisive effect on business processes, leading to increased costs and lower customer and employee satisfaction (Pipino et al., 2002, Redman, 1998). For example, data defects in pricing lead to inaccurate invoices; incorrect customer address data causes delivery delays and customer complaints; and incomplete accounting data produces a lot of manual effort. About 75% of organizations have identified costs originating from master data defects (Marsh, 2005). This is why companies are seeking to improve the quality of their data.

“Only what can be measured can be improved” is a phrase a number of authors have come up with (English, 1999, Wand and Wang, 1996, Wang and Strong, 1996). With regard to data quality, this statement suggests that in order to improve and evaluate data quality it has to be measured and monitored on a permanent basis. Many companies have run data quality initiatives focusing on actions to control and improve data quality (White et al., 2006). However, a comprehensive concept for data quality controlling is missing.

Controlling in terms of management is an instrument for performance measurement and correction in order to make sure that enterprise objectives are accomplished (Koontz, 1978). Applying this instrument to the domain of data quality management can help companies implement a sustainable mechanism for permanent monitoring and improvement of data quality. Therefore, the research objective of this paper is to identify success factors for implementing a data quality controlling system.

The research presented in this paper is case study research. That means the paper only investigates on the specific situation of the company the case study was conducted with. It does not present a generalizable approach for establishing data quality controlling.

The paper is structured as follows: Section 2 gives an overview of the state of the art in master data management, data quality management, and data quality controlling. Section 3 describes the research approach. The case under investigation is presented in Section 4. Section 5 discusses the findings from the case study. The paper closes with a conclusion and an outlook on further research in Section 6.

2. State of the art
Information systems provide data in a certain business context. When data is being used by human beings, they turn into information, and information finally turns into knowledge by interpreting and linking it for a given purpose (Boisot and Canals, 2004, Davenport and Prusak, 1998, Spiegler, 2000, Stahlknecht and Hasenkamp, 2005). Master data is a type of such information, describing a company’s essential business entities, such as suppliers, customers, products, employees, and assets (Dreibelbis et al., 2008, Loshin, 2008, Smith and McKeen, 2008). This data are used across multiple business processes. For example, customer data are used by order fulfillment as well as by accounting processes. In order to distinguish between
Master data and transactional data, different criteria are used, such as time reference, modification frequency, volume stability, or existential independence (Dreibelbis et al., 2008, Mertens, 2007, Newman and Logan, 2006).

Master data management (MDM) comprises all activities for creating, modifying or deleting master data classes, master data attributes, or master data objects (Smith and McKeen, 2008, White et al., 2006). As an application independent process, it ensures consistency and accuracy of data by providing a consistent understanding of master data entities, supported by mechanisms to use consistent master data throughout the organization and manage changes to it. These goals are achieved by implementing a corporate framework including three domains: organization, processes, and architecture (Dreibelbis et al., 2008).

Activities of MDM aim at providing master data of good quality (i.e. master data that is complete, accurate, timely, and well-structured) for being used in business processes (Karel, 2006, Loshin, 2008). Master data quality is defined by the degree of benefit perceived by a user using certain master data in a certain context. This is often referred to as “fitness for use” (Redman, 1996, Wang and Strong, 1996). The intended use is commonly described as a multi-dimensional construct consisting of a set of quality attributes, called data quality dimensions or metrics which are determined by the data consumer (Wang and Strong, 1996). Two of data quality metrics relevant to this paper are listed below following (Wand and Wang, 1996):

- **Accuracy**: The extent to which data are correctly representing an action or real world object.
- **Completeness**: The extent to which values are present in a data collection

Data quality management (DQM) comprises initiatives for improving the quality of data (Batini and Scannapieco, 2006). DQM as a part of MDM plays an important role for the success of MDM approaches.

The need for being able to ensure good data quality is particularly salient in decentralized organizations acting on a global level. Such companies typically are characterized by a heterogeneous landscape of data storing and processing systems due to a certain history of mergers and acquisitions, deviant requirements of business units, and different regulations in different countries (Ofner et al., 2009). Data quality (DQ) problems often occur when it comes to gathering information from a number of systems distributed across business functions or organizational boundaries (Batini and Scannapieco, 2006).

In order to be able to conduct data quality controlling that is both efficient and effective, general controlling requirements need to be taken into account. Table 1 shows these requirements together with questions relevant for controlling in the data quality domain.

<table>
<thead>
<tr>
<th>ID</th>
<th>Requirement</th>
<th>Source</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Definition of objectives</td>
<td>(Küpper, 1987)</td>
<td>What is needed for being able to quickly identify critical business data defects and take appropriate action?</td>
</tr>
<tr>
<td>R2</td>
<td>Organizational alignment</td>
<td>(Henderson and Venkatraman, 1993)</td>
<td>How should data governance organization and data quality controlling be related to each other?</td>
</tr>
<tr>
<td>R3</td>
<td>Planning and controlling</td>
<td>(Horváth, 2012)</td>
<td>What aspects are relevant for controlling and planning data quality in an enterprise?</td>
</tr>
</tbody>
</table>
How can the objectives of data quality controlling be integrated with the overall objectives of the enterprise and the more particular objectives of its employees?

What IT tools are needed to measure data quality?

At present, no DQM approach can be identified that addresses each of these requirements (see Table 2). The approaches listed in Table 2 are the most prominent ones that have occurred either in the research (Batini and Scannapieco, 2006, English, 1999, Pipino et al., 2002, Ryu et al., 2006, Wang, 1998) or the practitioners’ community (Bitterer, 2007, IBM, 2007, Thomas, 2006).

<table>
<thead>
<tr>
<th>Approach</th>
<th>Source</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total data quality management</td>
<td>(Wang, 1998)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Data Quality Management Maturity Model</td>
<td>(Ryu et al., 2006)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Data quality assessment</td>
<td>(Pipino et al., 2002)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Comprehensive methodology for data quality management</td>
<td>(Batini and Scannapieco, 2006)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Gartner's Data Quality Maturity Model</td>
<td>(Bitterer, 2007)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Total information quality management</td>
<td>(English, 1999)</td>
<td>○</td>
<td>○</td>
<td>●</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The IBM Data Governance Council Maturity Model</td>
<td>(IBM, 2007)</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The DGI Data Governance Framework</td>
<td>(Thomas, 2006)</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>○</td>
<td>●</td>
</tr>
</tbody>
</table>

Table 2: Analysis of DQM approaches regarding data quality controlling requirements

Legend: ○ not addressed ● partially addressed ● fully addressed

In his Total data quality management approach (Wang, 1998) has developed concepts, principles, and procedures for defining, measuring, analyzing and improving information products. The approach also includes a software instrument for information quality assessment. What the approach does not come up with is a clear organizational alignment, a broad information system as an instrument, and a concept for the integration of data quality objectives with performance management. The Data Quality Management Maturity Model by (Ryu et al., 2006) is a model that could be applied to evaluate and manage data quality in enterprises. However, this model targets the information system level only, and so it does not consider any organizational and performance management aspects. Data quality assessment (Pipino et al., 2002) and Comprehensive methodology for data quality management (Batini and Scannapieco,
are mainly focused on quality metrics for assessing data quality in enterprises. Gartner’s Data Quality Maturity Model (Bitterer, 2007) comes as a largely superficial, high-level maturity model for data quality management taking none of the above listed requirements fully into consideration. Total information quality management (English, 1999) provides deep insight into processes such as data cleansing and information quality assessment, but ignores the other aspects. And, finally, the IBM Data Governance Council Maturity Model (IBM, 2007) and the DGI Data Governance Framework (Thomas, 2006) have their focus only on data governance and organizational alignment with data quality objectives.

3. Research approach
3.1 Case study research design
Case study research is adequate if the phenomenon under investigation is still relatively unexplored (Yin, 2002). The paper uses a single-case study to focus on understanding the dynamics which are present within a certain setting (Eisenhardt and Graebner, 2007, Yin, 2002). (Yin, 2002) has proposed five components to be taken into account in case study research. The paper follows these components and takes them as a guideline for its research design. The research question (1st component) is derived from the (limited) knowledge of implementing concepts and systems for corporate data quality controlling. The central research question is:
How can companies implement a data quality controlling system in order to be able to monitor and improve data quality on a permanent basis?
Due to the limited knowledge with regard to the research question, the case study presented is to be considered as an exploration. In accordance with (Yin, 2002) regarding explorative case study research, it is unlikely that one can base one’s research on clear propositions (2nd component). However, a clear purpose is essential for each case study, and it should follow a set of criteria guiding the investigation. For this purpose, the paper uses a conceptual framework (Figure 1) based on (Küpper, 1987) as a guiding scheme for the investigation. The conceptual framework illustrates controlling as a system and assumes that controlling has a coordinating role between different sub-systems (Küpper, 1987).

Figure 1: Conceptual framework (Küpper, 1987)
The sub-systems depicted in Figure 1 have the following functions:

- **System of objectives** identifies DQM requirements from the company’s overall business objectives and derives DQ objectives.
- **Planning and controlling systems** define relevant DQ metrics (e.g. accuracy, completeness, etc.) and measurement processes.
- **Organization** determines the roles and responsibilities for DQ controlling and aligns them with corporate data governance.
- **Personal management systems** develop incentives and integrate DQ objectives with employees’ personal goals.
- **Information systems** support processes and methods for data quality controlling from an IT perspective.

The paper uses a single-case study design to understand how a controlling approach for data quality management is applied in a certain company. This means that the approach is implemented within the particular setting of the company investigated. Therefore, the unit of analysis (3rd component) sets the boundaries with regard to generalizability of the results.

The conceptual framework functions as the logic which links the data to the propositions (4th component) and the instrument structuring the information and findings gained from the case study (5th component).

### 3.2 Case selection and data collection

As part of the study (Otto and Ebner, 2010) a survey was conducted comprising 40 experts from different organizations. The aim of this survey was to identify data quality controlling mechanisms. As the results of the survey suggest, organizations measuring data quality mainly do so by measuring data quality metrics at irregular intervals and by measuring the quality of data within a certain database. Only few organizations measure DQ on a permanent basis. The survey also shows that organizations which have implemented a comprehensive DQ controlling system are very rare.

The case presented in this paper refers to a manufacturer of power tools for the global construction industry. At the time of the case study, the company had already successfully established a comprehensive data quality controlling system for its customer data.

Data collection was done by conducting three expert interviews. An overview of the interviews conducted is given in Table 3. Each interview was recorded in writing and then transcribed, analyzed and evaluated afterwards. The interviews have been conducted with the same person.

<table>
<thead>
<tr>
<th>Interview No.</th>
<th>Organizational function of interviewee</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview 1</td>
<td>Global process manager and project leader for customer data quality within central DQ unit</td>
<td>120 min.</td>
</tr>
<tr>
<td>Interview 2</td>
<td></td>
<td>180 min.</td>
</tr>
<tr>
<td>Interview 3</td>
<td></td>
<td>120 min.</td>
</tr>
</tbody>
</table>

**Table 3**: Interview overview

A cost-benefit analysis for the case has not been conducted yet because the corporate strategy determines the justification for DQM directly.
4. Case Presentation

4.1 Company profile
The case presented is about one of the largest manufacturers of power tools worldwide, hereinafter called MPT. While MPT was founded as a family business in (1941), it today has about 22,000 employees working in 120 countries. In 2011, sales revenues were approximately 4 billion Swiss Francs. The company is focused on developing, manufacturing, selling and distributing high-quality products, such as equipment and systems for covering, drilling and demolition. The company’s main customer group consists of professionals working in the construction and building maintenance industry. As MPT offers its customers five different sales channels, two-thirds of the employees work directly for the customer in sales organizations and engineering, which means a total of more than 200,000 customer contacts every day.

4.2 Business strategy
MPT has defined a strategic “Vision 2015” in order to ensure profitable and sustainable growth. The strategy is focused on customer satisfaction as one of the main drivers of the business. Being in direct contact with customers and conducting effective customer relationship management (CRM) have been identified as two main pillars for achieving a high level of customer satisfaction. For this purpose it is essential to know as much as possible about customers. In striving for this essential goal the company management has recognized that it is of crucial importance that the quality of its customer master data does not drop below a certain level, making a comprehensive and reliable data quality controlling system indispensable.

4.3 Controlling customer data quality
The company’s business strategy has been the main driver of the implementation of a comprehensive controlling system for customer data quality. As mentioned above, MPT provides customers with services and products using five different sales channels. The challenge here is to keep the customer data consistent across all channels, as high-quality data is considered to ensure a smooth flow of different customer processes (such as delivery, billing, etc.) and as the company highly benefits from efficient and effective lead and customer relationship management. That is why MPT has implemented a comprehensive customer data quality controlling system in order to be able to monitor and improve customer DQ on a permanent basis. Using the conceptual framework presented above, the sub-sections outline how MPT designed and implemented the sub-systems of the comprehensive controlling system.

4.3.1 System of objectives
In order to be able to achieve the company’s overall business objectives, the following DQ objectives were identified as critical for implementing customer data quality controlling (these DQ objectives constituted the basis for the definition of tasks in other controlling related sub-systems):
- create transparency with regard to the customer DQ situation
- control and monitor customer DQ trends on a permanent basis
- identify need for action
- launch appropriate data cleansing and quality improvement activities
4.3.2 Planning and controlling systems

MPT was supposed to identify and implement a single index to measure DQ in the customer data domain and allow for comparison: the Data Quality Index (DQI). An arbitrary DQI value of 90% was set to be achieved in each country. The DQI was based on two relevant quality metrics: accuracy and completeness. These metrics have been identified as relevant for customer processes such as billing and delivery and are built on a set of business rules for critical customer data attributes.

In order to follow the business rules and achieve the required DQI value, the planning and controlling processes were divided into two areas:

- **Proactive customer DQM**: The customer data creation process has been adjusted so that certain checks based on the business rules take place during data entry already. This way the risk of entering corrupt data is minimized from the outset as far as possible.
- **Reactive customer DQM**: To improve the quality of existing customer data, a number of activities have been defined, such as identifying data defects through permanent data monitoring. The responsible corporate function triggers required data cleansing activities for affected customer data attributes.

4.3.3 Organization

MPT defined an organizational process comprising a number of roles for customer master data quality controlling. Country managers are the top responsible function for customer DQ. Country managers delegate customer DQ related tasks to local process experts (LPEs). LPEs have the local responsibility for customer DQ and coordinate data cleansing and DQ improvement activities on country level. Depending on the type of data fields, they work with different local departments (such as sales, supply chain, or finance) and trigger the data correction process. Communication between LPEs and departments is bidirectional. That means the departments are also able to trigger data correction processes if they discover any problems in their processes which are likely to be due to poor data quality. In some cases (mass changes, for example) they are also allowed to instruct the LPE to execute changes.

![Organizational process diagram](image)

**Figure 2**: Organizational process
Figure 2 shows the organizational process for correcting data defects and reporting on the status of DQI. The global process manager (GPM) is responsible for the quality of customer data on a global level and is supposed to pursue DQ related issues across all branches. The GPM can report issues related to customer DQ directly to the chief of marketing on a corporate level, who can escalate the problems directly to the respective country manager or regional manager. The GPM monitors the DQ level of each country on a monthly basis.

4.3.4 Personal management systems
Regional managers specify DQ objectives for country managers and partly integrate these objectives into the country managers’ personal objectives. This leads to similar objectives for LPEs, who report directly to the country manager and are active on the operational level. Additionally, in many countries certain financial incentives (rewards for LPEs and country managers, for example) are fixed for achieving a certain level of DQ.

4.3.5 Information systems
Information systems support the organization in implementing defined processes and responsibilities both on a local and a global level. At MPT, a SAP ERP system has been defined as the single source of truth, containing the relevant customer data. In order to ensure high data quality on this platform, different tools have been developed and are being used:

- The DQ monitoring tool supports reactive DQM. In SAP a monthly job is scheduled which exports all critical customer data fields for each country into a flat-file and sends it to the respective LPE via e-mail. The DQ monitoring tool is an access based report which helps LPEs analyze the actual DQ based on the extracted flat-file. LPEs import the file into the DQ monitoring tool and the tool identifies all critical data defects based on defined business rules and shows the current DQI of the country.

- The mass maintenance tool is an in-house development based on SAP ERP, allowing to make changes and corrections to large data volumes in a short time. LPEs are the main users of this tool, making corrections by order of different business departments. The tool also supports reactive DQM.

- The data creation workflow is a workflow for the customer data creation process supporting proactive DQM by facilitating various approval steps. The main user of this tool is the customer interaction center, which is responsible for data creation. In cooperation with external data providers certain data quality checks can be conducted within the workflow processes, avoiding creation of corrupt data.

- A mobile application for data correction, developed by MPT, helps sales representatives capture the correct data on site and trigger the correction processes to be implemented by a workflow. The mobile application has access to a set of predefined data fields and supports reactive DQM.

- The data correction workflow is a mobile application that triggers the data quality workflow which is in use by the back office responsible for customer data. It checks the plausibility of changes and approves or rejects them. This workflow supports both proactive and reactive DQM.

4.4 Summary
The case of MPT shows that for implementing a comprehensive controlling system for data quality different aspects known from controlling can be taken into account.
MPT implemented all controlling sub-systems as defined by the conceptual framework (Küpper, 1987). The system of objectives contains concrete data quality goals which were derived directly from corporate objectives and the corporate strategy. The definition of a data quality index based on identified quality metrics constitues the basis of the planning and controlling system at MPT. Reactive and proactive customer data quality management with clearly defined processes makes it possible to define clear organizational responsibilities and activities based on data defects that have occurred. Furthermore, MPT defined a single role to be responsible for the quality of its data. This role has to coordinate all necessary activities to improve data quality and achieve the quality level required. Integration of data quality objectives into personal goals provides incentives for local process experts as well as country and regional managers to invest more in the quality of customer data.

The most important support tools for MPT for implementing the controlling system are the IT applications which have been developed internally. These applications make it possible to monitor data quality within defined periods of time and initiate required data correction activities.

According to the statement of interviewee from MPT, the quality of customer data has increased significantly since the controlling approach has been implemented. This has had also a positive impact on customer satisfaction, which several customer surveys have shown. Another indication for the improvement of customer data quality is the decreasing number of wrong deliveries and customer complains caused by incorrect address data.

At the initial state of the measurement process the average of each DQI was about 50%. After approx. 3 years executing regular measurements the average is about 85%. This is also an indicator for improving customer data quality.

As the case study shows, applying a controlling approach for data quality management can be very useful and can bring noticeable benefits for companies.

6. Case Discussion
Whereas many organizations and DQM related literature sources mainly see data quality controlling activities in defining metrics and measuring the quality for identified metrics, the case presented in this paper shows another facet of data quality controlling. In literature on DQM no approach can be identified that deals with a DQ controlling concept. Such a concept requires technical as well as organizational and strategic aspects to be considered. The case presented in this paper shows that such a controlling concept is important for enterprises to control data quality on a permanent basis.

The following success factors can be considered key findings from the case studied:

- Controlling as a tool: The most important success factor is to consider controlling as a tool for permanent data quality management. Controlling concepts can be transferred to the data quality domain to ensure better data quality.
- Clear mandate from management: Sponsorship by top management can be seen as key for implementing DQM controlling. It is indispensable for establishing strategic and organizational structures.
- Definition of DQ objectives: Defining organizational DQ objectives is important for the implementation of a successful controlling concept
- Strong IT support: Appropriate IT tools are essential to support the entire controlling concept for data quality.
7. Conclusion and further research

Based on the findings from a single-case study the paper describes the process of implementing a comprehensive controlling system for DQM. The system comprises all relevant controlling sub-systems as defined by (Küpper, 1987) and takes into account all the business engineering layers (strategy, organization, processes, and information systems) as defined by (Österle and Blessing, 2003). The paper makes a contribution to the advancement of the scientific body of knowledge by showing a possible implementation of controlling concepts in the data quality domain that can help companies monitor and improve their DQ activities. The case study suggests that a successful implementation of a system to control DQ is supported by a corporate approach.

Only few organizations have implemented a controlling system for DQM so far. Therefore, it is important to identify the essential tasks to successfully introduce such a system. Against the background of the case investigated, future research will focus on gaining more insight into activities on controlling master data quality. To do so, further case studies will be conducted that may build the foundation for identifying best practices and allow for generalization of results. A framework containing all necessary tasks to implement a DQM controlling system would be a desirable outcome of this research-in-progress. This framework should allow assessing the complete range of DQ problems from their root causes to their impact. Also, a sample evaluation with a focus group is planned to verify the findings.
References


