Health Informatics Meets eHealth
D. Hayn and G. Schreier (Eds.)

doi:10.3233/978-1-61499-759-7-259

Data Quality- and Master Data Management – A Hospital Case

Klaus ARTHOFER[a,1], Dominic GIRARDI[b]
[a]*Degree Course Process Management in Health Care; Upper Austria University of Applied Sciences, Steyr, Austria*
[b]*Research Unit Medical-Informatics, RISC Software GmbH, Hagenberg, Austria*

Abstract. Poor data quality prevents the analysis of data for decisions which are critical for business. It also has a negative impact on business processes. Nevertheless the maturity level of data quality- and master data management is still insufficient in many organizations nowadays. This article discusses the corresponding maturity of companies and a management cycle integrating data quality- and master data management in a case dealing with benchmarking in hospitals. In conclusion if data quality and master data are not properly managed, structured data should not be acquired in the first place due to the added expense and complexity.

Keywords. Data management, Data quality, Master data management, Data governance, Hospital

1. Introduction

Hospital costs are increasing steadily, making further efforts to control costs both inevitable and essential. Simultaneously, there is a growing need for even better medical quality for the treatment of patients. LeiVMed, a non-profit programme provided by the degree course Process Management in Health Care (in German: Prozessmanagement Gesundheit, PMG) at the University of Applied Sciences Upper Austria's Steyr Campus, has been pursuing this mission for about eight years. "LeiVMed" stands for "Leistungsvergleich Medizin" (in English: benchmarking in health care) and is used to prepare administrative and medical data to provide information about treatment specific surgical patient outcomes, variable treatment costs, processes of care, and data for guiding local quality improvement efforts as well as cost efficiency programs. PMG developed LeiVMed over the course of several practical and research projects. Besides providing support for hospitals it serves scientific purposes while also creating synergies with teaching.

LeiVMed currently analyzes the treatment processes of hip total endoprothesis, prostatectomy, hernia, strumectomy, cholecystectomy, colon operation and rectum operation for three Austrian hospital operators. The data is acquired from hospital data bases as well extracted from unstructured medical documents by trained data nurses. LeiVMed analyzed so far about 15 000 medical cases of the treatment process types mentioned above plus appendectomy (appendectomy cases are no more analyzed, because they are largely acute cases inadequate to interpret). "Medical cases" in LeiVMed imply cases from the perspective of doctors and managers and not billing cases

[1] Corresponding Author: Dr. Klaus Arthofer, FH OÖ Studienbetriebs GmbH, Wehrgrabengasse 1-3, 4400 Steyr/Austria, E-Mail: Klaus.Arthofer@fh-steyr.at

being subject of hospital applications. LeiVMed's medical cases thus involve mostly at least one stationary and one ambulatory billing case including the corresponding medical procedures. Besides this LeiVMed's medical cases comprise anonymized personal data and medical risk factors of the according patients, medical complications and laboratory values.

Regarding cost indicators LeiVMed especially addresses the standardization of processes. In respect of standardization LeiVMed shows the depreciation of the medical treatment processes mentioned above in medical departments from the corresponding evidence based medical guidelines. These cost and process indicators are complemented with medical outcome indicators represented by medical complication ratios. Medical complications are of central relevance for LeiVMed's hospital customers and have to be manually extracted from unstructured medical documents by study nurses. Data quality and master data management in LeiVMed focus on internal procedure data (internal from the perspective of a hospital), because the entry of internal procedure data in particular is generally still inadequately organized in hospitals. Although LeiVMed is limited to specific (but frequent) treatment processes and focuses on internal procedures of especially radiology and laboratory departments, it takes about 1000 internal procedure concepts into account. LeiVMed uses SNOMED CT (Systematized Nomenclature of Human and Veterinary Medicine) as a standard for its procedure definitions and organizes the matching of the proprietary procedure definitions of its hospital customers with the SNOMED definitions as well as its remaining concept definitions based on an ontology.

LeiVMed's major success factors are fairness and accuracy. Fairness is addressed with analytical models, specifically risk adjustment ensuring indicators related to the special characteristics of the patient sample analyzed. Accuracy is primarily based on data quality. The management of data quality- and master data management in LeiVMed as well as in hospitals will be discussed in this article. Data quality- and master data management is not only a success factor for LeiVMed, it is of particular interest for all organizations using database applications. This is because data quality- and master data management is renowned (garbage in, garbage out), and it is more of an organizational issue and perhaps for this very reason remains highly problematic.

2. Methods

This article discusses data quality- and master data management in the case of LeiVMed. Both authors of this article have been working on LeiVMed since its inception, have been publishing related articles, especially on ontology-based data analysis, and worked on several similar projects dealing with sectors other than healthcare. Therefore and due to both the general relevance and the general validity of the topic, this article basically addresses data quality- and master data management for all sectors and shows the corresponding implications via LeiVMed.

This article starts with an integrated definition of data quality- and master data management followed by a literature review on data quality and master data management in companies along with their maturity over the last decade. The literature review used scientific literature databases such as SpringerLink, IEEExlore and SciVerse ScienceDirect, along with queries in Google Scholar. A survey taking the form of semi-structured, oral interviews with data management experts at six large Austrian hospital operators supplements the literature review regarding hospitals.

This discussion on the maturity of companies regarding data quality- and master data management results in the approach for data quality- and master data management of LeiVMed.

3. Data quality and master data management

Master data is independent of transaction data and is referenced in transaction data. Customer and product data are examples of master data. Order data referencing customer and product data are transaction data. "Master data management is a business function that encompasses all planning, monitoring and provisioning activities associated with master data, in order to ensure master data quality and thus its suitability." [13] In contrast to master data management, data quality management focuses on the processes, methods and tools for analyzing and cleansing data.

Transaction data quality and master data quality are often seen as "two sides of the same coin". If transaction data quality is poor, master data quality receives less attention. If the master data is poorly maintained, the quality of the transaction data is necessarily limited [10], [13]. Therefore on the one hand data quality as well as master data have to be managed autonomously and their management has to be integrated, on the other hand any subordination of one area to the other is inappropriate. Particularly preventive data quality activities (e.g. data governance) influence a company's master data management as they define responsibilities for data maintenance. Hence, preventive data quality management does have an effect on a company's master data management [10].

4. Maturity of data quality and master data management

Dippold et. al analyzed in 2005 that corporate data management is generally about twenty years behind the current state of research. Process organization and to a certain extent human resources are in particular need of improvement with regards to the management of data quality and master data. Important data is better managed and unimportant data tends to be neglected. As mentioned above, industries differ in terms of their maturity in managing data quality and master data [2]. Otto cites a study from BARC dating from 2008, stating that two thirds of companies view their master data management as immature or not fully mature and concludes that the majority of companies are in the midst of organizing or reorganizing master data management [12].

In a survey based on partially structured interviews with data management specialists at six large Austrian hospital operators in 2012 consensus has been found regarding the importance of data quality and master data management (even in areas irrelevant to billing). However, no expert claimed to regularly examine the quality of relevant non-billing data. Specific departments were responsible for defining master data depending on the master data type (e.g. procedure master data by controlling department), with this generally being agreed within departments at an individual site and not throughout the holding. No hospital operator allocated such responsibilities to specific persons. The maintenance of master data in the applications in general was primarily the work of the IT department, particularly if doing so required advanced computer skills. No hospital operator offered courses on master data or data entry rules. The relevant users simply received an e-mail. On the whole, both the data quality and master data management structures and processes were implicit rather than explicit. As the focus of

the experts' comments thus fell on IT, data management appears to both IT staff as well as staff in other departments as more a technical issue [14]. Therefore only a limited level of maturity in managing data quality and master data can generally be attributed to these hospital operators. Probably as the losses made by austrian, public hospitals are compensated with operational loss coverage, data relating to internal procedures, surgical staff functions etc., however cost-relevant it may be, is regarded as less important. Thus these hospitals themselves also focus on important data (from their perspective) when it is a question of data quality and master data management.

In the literature, data management in general is viewed at the strategic, organizational and systemic levels [3], [12], [16], [4]. The general focus tends to be on the systemic level – especially if the literature does not explicitly focus on data management. CobiT (Control Objectives for Information and Related Technology), the international IT governance framework, also examines data management in its "deliver and support" domain at the systemic level and not in the "plan and organize" domain. Although CobiT does not want to focus on how IT requirements are implemented but on which IT requirements need to be implemented [7]. Thus data management generally seems to be seen as more of a technical domain, what implicates its potential at both the strategic and the organizational level.

A recent study published in January 2016 on the state of enterprise data quality from 451 Research commissioned by Blazent, a US-based company focusing on data management, confirms this hypothesis. They identified a “disconnect between responsibility and accountability for data quality”. IT departments are mainly responsible for data quality management and managerial teams are ultimately (meaning “not explicitly”) held accountable for it. But these parties are poorly aligned with each other, which of course has negative impacts on data quality. The authors of the study believe that this finding explains the contradiction between the other two key findings (besides the latter) of this study: on the one hand respondents generally have doubts about the effectiveness of the data quality management initiatives in their companies, on the other hand respondents believe, that a substantial portion of business value can be lost due to poor data quality. Therefore they conclude that there is a “laissez-faire attitude toward the quality of data and the DQM practices in their organizations” and consider it “to be exacerbated due to the anticipated growth of data and plans for future projects that drive data creation and therefore need for quality management” [34]. A study from Omikron in 2013 surveying 200 business managers from the German speaking area in Europe came to very similar conclusions: neither business departments nor IT departments promote data quality, but both know, that “business nowadays is based on data and not on computers”. They explain this finding mainly with the insufficient cooperation between business departments concerning data quality [18].

In summary a certain progress in data quality management has been achieved over the last decade. Unfortunately this progress seems to be too weak to keep pace with the anticipated relevance of data. Regarding the evolution of the market for master data management software [5], the complex, ever-increasing and constantly changing regulatory environment [15] and last but not least the evolution of data management assessment tools (e.g. CMMI DMM) [1] the evolution of master data management may be comparable with that of data quality management.

5. LeiVMed's approach to data quality and master data management

In their survey from 2011 Haug and Arlbjørn found that the most significant barriers to master data quality are:

- Poor delegation of responsibilities for maintaining master data (which is the most important barrier)
- The absence of master data monitoring routines
- The lack of employee skills [6]

Thus the task of managing data quality and master data needs to be assigned to a person or group of persons as specifically as possible. The same person or group of persons needs to perform regular and successively expanded tests of data quality and master data maintenance to allow employees to see the data quality problems they are causing at regular intervals. Employees can then learn how to improve and thus be motivated to consistently generate data with sufficient quality.

LeiVMed supports its hospital customers in defining master data, particularly internal procedure concepts, and training staff to manage relevant master data along with the semantic testing of the transaction data (in cases of pneumonia thorax X-rays for example need to be recorded). As LeiVMed thus takes over a large portion of the quality and master data management of the relevant data, the hospital customers have sourced out aspects of managing data quality and master data besides medical controlling activities. This reduces the problem of assigning roles as well as that posed by the lack of master data monitoring routines and employee skills substantially. This in turn enables the hospital customers to adapt their own resources whilst already taking advantage of their medical benchmarking activities.

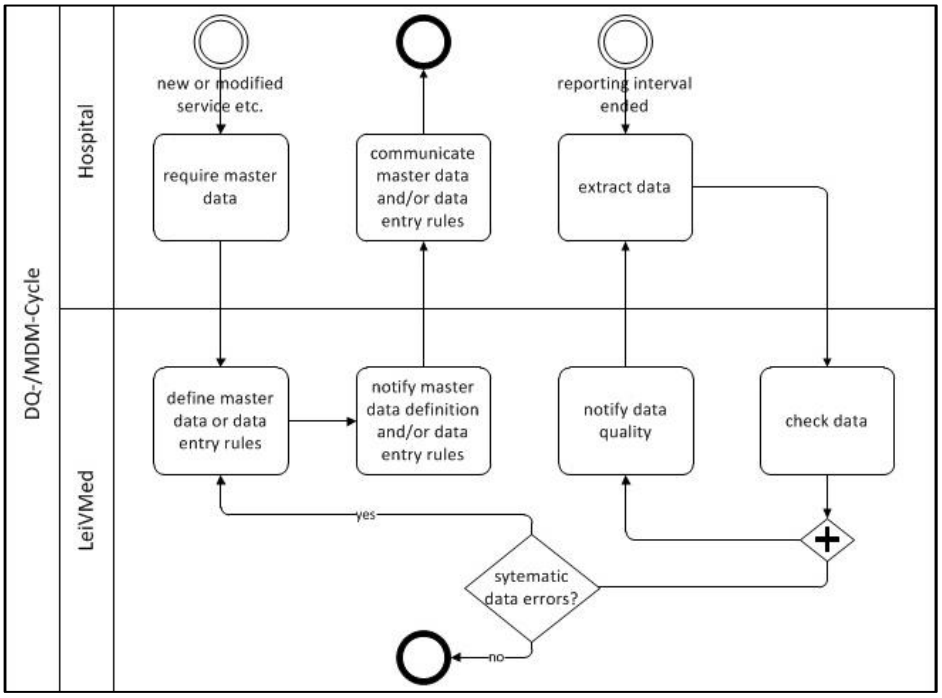

Figure 1. Data quality and master data management cycle

The data quality and master data management process is presented (simplified) as a cycle in Figure 1, also indicating the potential for continuous improvement of transaction and master data quality. LeiVMed calculates benchmarks on a quarterly basis. It is sensible to institutionalize testing of the data before loading it in a data warehouse. The data quality check is not merely used in the immediate data correction for data analysis but also needs to be integrated in the organization to successively improve data quality. Thus data quality in the online transaction systems is improved fostering their support for the execution of business processes. Inferior master data means that often only a fraction of the functions from applications introduced with considerable effort and cost can then be utilized [17]. After all, the effort needed to correct data errors will continue to decrease and the quality of the data assessed will improve, making data analysis more informative.

Both syntactic and semantic checks need to be performed on the extracted data. Adequate plausibility tests need to be defined along with the use of software tools with rules engines, fuzzy search methods etc. which are complex and time-consuming. The outsourcing of data assessment and testing also involves data confidentiality issues, which can be dealt with by using encryption and pseudonymization as well as non-disclosure agreements etc.

After checking the data, LeiVMed will notify the hospital about the data quality. If systematic errors are found in the data check, eg a medical consilium of a certain department is not recorded in general, they are often caused by poor master data definitions or data entry rules, or the application users failing to properly utilize such definitions and rules. This situation constitutes the link between the management of data quality and management of master data in Figure 1. Data quality management thus extends into master data management for systematic errors, whereby LeiVMed edits or explicitly sets master data definitions and/or data entry rules. This is in keeping with the initial definition of master data and data entry rules following a specific request from the hospital.

This approach combining outsourcing aspects of data quality (from the perspective of the hospitals) with master data management has led to the following benefits at LeiVMed's hospital customers:

- Surmounting the most significant barriers to master data quality (which are also applicable for the quality of transaction data):
  - Explicitly defined persons in charge of core activities in data quality and master data management
  - Provision of master data monitoring routines
  - Provision of data quality and master data skills
- Capability to analyze data on a scientific level
- Potential for continuous improvement of the maturity of the hospital's data quality and master data management and support in the development of an appropriate attitude

This approach initializes only with great engagement on the part of both LeiVMed and the hospital. The data quality and master data management activities have to be coordinated between LeiVMed and the hospital and some resistance to organizational change in the hospital has to be overcome. Finally the potential for improvements can only be realized with sustained adequate leadership in the hospital.

6. Discussion

Management of data quality and master data are crucial for adequate data analysis and business decision-making as well as to the smooth running of business processes. Given the lack of maturity in this regard, the management of data quality and master data offers huge potential, not just in hospitals but for organizations in general.

On top of the methodical and technical hurdles, the assignment of responsibility for data quality and master data management to one or more persons presents a special challenge. At the same time the strategic and organizational perspective (see Chapter 4) need to be taken into account. Data management is generally seen as an IT issue but the IT department does often not really understand the precise semantics of a considerable amount of data. This creates a vacuum in terms of responsibility, as the operative departments view data management as an IT issue, but IT cannot work properly without input from the departments. In addition, attention needs to be paid to the associations between the activities of management of data quality and management of master data as well as to their triggers.

Data quality management and master data management are important to the users of all business applications and not merely data analysis, decision-making and other applications reliant on adequate data. Doctors in hospitals have long complained about the excessive effort involved in entering data (e.g. encoding of procedures). Adequate management of data quality and master data would not only provide for more meaningful analyses but probably also demonstrate that a good deal of data is input to no real purpose.

The importance of data quality and master data management may not be ignored. If data cannot be properly managed or if there is no will to do so, the data should not be acquired, structured or even encoded in the first place due to the added expense (recording effort, cost of IT solutions) and complexity of doing so.

References

[1] CMMI Institute, 5 Steps to Jump-start Your Data Management Program, available on: http://cmmiinstitute.com/sites/default/files/resource_asset/Data%20Management%20Maturity%20DMM%20ebook.pdf

[2] Dippold R, Meier A, Schnider W, Schwinn K, Unternehmensweites Datenmanagement. Von der Datenbankadministration bis zum Informationsmanagement, Vieweg&Sohn, 2005, pp 175-182

[3] Dippold R, Meier A, Schnider W, Schwinn K, Unternehmensweites Datenmanagement. Von der Datenbankadministration bis zum Informationsmanagement, Vieweg&Sohn, 2005, p 12

[4] Gansor T, Scheuch R, Ziller C, Master data management. Strategie, Organisation, Architektur, dpunkt, 2012, p 68

[5] Goetz M, The Forrester Wave™: Master Data Management, Q1 2016. The 12 MDM Providers That Matter Most And How They Stack Up, available on: http://www.forrester.com/pimages/rws/reprints/document/119980/oid/1-ZFXE6A

[6] Haug A, Arlbjørn J, Barriers to master data quality, Journal of Enterprise Information Management, vol 24, 2011, pp 288-303.

[7] KPMG Austria, CobiT 4.0. Deutsche Ausgabe, 2005, pp 157 – 159, available on: http://www.isaca.at

[8] Lehmann C, Roy K, Winter B, The State of Enterprise Data Quality: 2016. Perception, Reality an the Future of DQM, 451 Research, 2016

[9] Messerschmidt M, J. Stüben, S. Gehrls, Lean Master Data Management - agieren statt reagieren, Information Management und Consulting, vol 25, 2010, p 75.

[10] MesserschmidtM , Stüben J, Verborgene Schätze - Eine internationale Studie zum Master-Data-Management, PwC, 2011, pp 76-77.

[11] Otto B, Hüner K, Österle H, Toward a functional reference model for master data quality management, Inf Syst E-Bus Manage (2012) 10, p 399

[12] Otto B, Stammdatenmanagement: Datenqualität für Geschäftsprozesse, Praxis der Wirtschafsinformatik HMD 279, 2011, pp 5-16.
[13] Otto B, Weber K, Data Governance, In: K. Hildebrand et. al., Daten- und Informationsqualität, Bd. 4, Vieweg+Teuber, 2011, pp 235-238.
[14] Pointner C, Datenqualitäts- und Stammdatenmanagement in oberösterreichischen Krankenhäusern,. bachelor thesis at University of Applied Sciences Upper Austria, 2012.
[15] Sadiz S, Governatori G., Managing Regulatory Compliance in Business Processes, In: J. Brocke, M. Rosemann, Handbook on Business Process Management 2, 2nd edition, Springer, 2015
[16] Schemm, Zwischenbetriebliches Stammdatenmanagement. Lösungen für die Datensynchronisation zwischen Handel- und Konsumgüterindustrie, Springer, 2009, p 9
[17] Scherer E, Problemfeld Stammdatenpflege, Monitor Special 12A, Bohmann, 2005
[18] Vaske H, Studie über Datenqualität. Datenqualität lässt in vielen Firmen zu wünschen übrig, Computerwoche, 2013, available on: http://www.computerwoche.de/a/datenqualitaet-laesst-in-vielen-firmen-zu-wuenschen-uebrig,2541967
[19] Wand Y, Wang R, Anchoring data quality dimensions in ontological foundations,vol 39, 1996, pp 86-95