University of Passau

Area of Study: Business Administration

Data Quality Management in Corporate Practice

Master Thesis

Supervisor

Prof. Dr. Franz Lehner

SS 2020

cand. M.Sc.

Simon Rodler, B.Sc.

89123

Münchner Straße 29f

82041 Oberhaching

simon@rodlers.de

Passau, *31.08.2020*

Abstract

The 21st century is characterized by a rising quantity and importance of Data and Information. Companies utilize these in order to gain and maintain competitive advantages. Therefore, the Data and Information is required both in high quantity as well as quality. But while the amount of Data collected is steadily increasing, this does not necessarily mean the same is true for Data Quality. In order to assure high Data Quality, the concept of Data Quality Management (DQM) has been established, incorporating such elements as the assessment of Data Quality as well as its improvement. In order to discuss the issue of Data Quality Management, this paper pursues the following goals:

(1) Systematic literature search for publications regarding Data Quality Management (Scientific contributions, Practice reports etc.)
(2) Provision of a structured overview of the identified references and the research material
(3) Analysis and evaluation of the scientific contributions with regards to methodology and theoretical foundation
(4) Current expression of DQM in practice, differentiated by organization type and industry (based upon the entire research material) as well as assessment of the situation (how well are the design recommendations based upon research results)
(5) Summary of unresolved issues and challenges, based upon the research material

Keywords: Data Quality Management, Systematic literature review, Data Quality, Information Quality, Data Quality Management Research

Abstrakt

Das 21. Jahrhundert ist geprägt durch eine steigende Quantität und Wichtigkeit von Daten und Informationen. Firmen nutzen diese, um Wettbewerbsvorteile zu erlangen und auszubauen. Aus diesem Grund sind Daten und Informationen sowohl in großer Quantität als auch in hoher Qualität notwendig. Während jedoch die Quantität gesammelter Daten stetig zunimmt, gilt dies nicht notwendigerweise auch für die Datenqualität. Um eine hohe Datenqualität zu sichern, wurde deswegen das Konzept des Datenqualitätsmanagements etabliert, welches Aufgaben wie die Bewertung der Datenqualität als auch deren Optimierung beinhaltet. Diese Arbeit verfolgt zur Betrachtung des Datenqualitätsmanagements folgende Ziele:

(1) Systematische Literatursuche nach Veröffentlichungen zum Thema (wissenschaftliche Studien, Praxisberichte usw.)
(2) Strukturierte Darstellung der identifizierten Quellen und des Recherchematerials im Überblick
(3) Analyse und Bewertung der wissenschaftlichen Rechercheergebnisse in Bezug auf die verwendete Methodik und die theoretische Fundierung
(4) Aktuelle Gestaltung bzw. Ausprägung des DQM differenziert nach unterschiedlichen Unternehmenstypen und Branchen (aufbauend auf dem gesamten Recherchematerial) und Bewertung der Situation (d.h. wie gut sind z.B. Gestaltungsempfehlungen wissenschaftlich abgesichert?)
(5) Zusammenfassung der offenen Fragen und Herausforderungen auf Basis des Recherchematerials

Schlüsselbegriffe: Datenqualitätsmanagement, Systematisches Literaturreview, Datenqualität, Informationsqualität, Datenqualitätsmanagementforschung

Table of Contents

Table of Figures

List of Tables

UNIVERSITÄT PASSAU

Table of Appendices

List of abbreviations

DQM: Data Quality Management

TDQM: Total Data Quality Management

CDQM: Corporate Data Quality Management

CC DQM: Competence Center Data Quality Management

IQM: Information Quality Management

DQ: Data Quality

IQ: Information Quality

IP: Information Product

1. Introduction

Data and its usage increasingly influence both individual and organizational everyday life. It is very heterogenous, from analyzing and predicting the movements of markets to the monitoring of infection chains. All applications however require a sufficient amount of Data of an adequate quality in order to serve their purpose appropriately. Therefore, the assurance of Data Quality is of vital importance for digital enterprises, a task which is referred to as Data Quality Management (DQM). Data Quality Management as a part of the Data Management function of an enterprise aims to assure that its Data meets the necessary requirements in order to support the enterprises business operations.

Data Quality and its management, however, is a diverse and extensive field, with many different views on what Data refers to, how its quality can be expressed and assessed and how Data Quality can be managed within enterprises. Therefore, there is a need for a structured approach towards this issue and the research conducted with regards to it. Before a literature review regarding DQM can be conducted the concepts of Data, Data Quality and its Management have to be introduced in order to foster some intuitive understanding. Besides, a short introduction is given to the means of how DQM can be organized and managed, Data Governance (Österle & Otto, 2016).

1.1 Data

Data and Information are concepts often referred to in the context of digitalization and Big Data, however there is no absolute differentiation between the two (Österle & Otto, 2016). According to ISO/IEC 2382-1, Data is the formalized representation of attributes of business objects, while sometimes also being regarded as constituting Information (Österle & Otto, 2016). Another view is that Data refers to the technical aspects, while Information refers to non-technical aspects (Zhu et al, 2014). However, due to a lack of clear and comprehensive differentiation in the reviewed literature, this paper adopts the concepts of Data and Information synonymously (Zhu et al, 2014; Österle & Otto, 2016). According to Österle & Otto (2016), there are also different levels of aggregation with regards to Data, five in total:

(1) Elements: Data elements form the first level of aggregation and the base elements for all Data.

(2) Records: Data records form the second level of aggregation, which can be seen as the instantiation of Data objects such as customer records

(3) Tables: Tables form the third level of Data aggregation according to Österle & Otto (2016), containing several Data records.

(4) Databases: Databases form the fourth level and contain several tables (Österle & Otto, 2016). They are the primary storage mechanism for Data

(5) Resource: Data resource forms the fifth and final level of Data aggregation according to Österle & Otto (2016). It contains the entirety of the Databases of an enterprise.

Another aspect of Data according to Österle & Otto (2016) is Enterprise Master Data or Master Data, which following the ISO 8000 standard is Data unique to the organization. Master Data can both be global or local and are more static than other Data, as well as more constant in terms of volume. It includes Data records such as customers or employees (Österle & Otto, 2016). In difference to Master Data, Transactional Data varies in volume and is subject to frequent changes.

Besides Master Data and Transactional Data, another often mentioned type of Data is Metadata (Sautter et al, 2018), which is Data referring to other Data. A distinction can be made between three types of Metadata, those being (1) descriptive Metadata used for identification and (2) structural Metadata, referring to structure, attributes, and versions of the Data. Finally, (3) there is administrative Metadata which refers to both technical and methodological aspects of Data creation and access rights (Sautter et al, 2018).

1.2 Data Quality

Data Quality (DQ) is a multidimensional concept, consisting of usually five overall dimensions according to Österle & Otto (2016). These five dimensions are:

(1) Correctness: This dimension of DQ refers to the need of Data to correctly represent the attributes of the real-world object it represents.

(2) Consistency: This dimension refers to the need for different representations of the same real-world object to be consistent with each other.

(3) Completeness: This dimension of DQ refers to the need for all attributes of the real-world object to be represented by the Data.

(4) Timeliness: This dimension refers to the need for the Data to always represent the real-world object and its attributes, at any point in time.

(5) Availability: This last dimensions of DQ according to Österle & Otto (2016) refers to the need to be able to access the Data when it is needed.

However, besides these five common DQ dimensions, there is a great number of additional dimensions used in the literature. Strong et al (1997) differentiate in four categories of DQ dimensions in order to categorize these. The first of these categories, intrinsic DQ, contains such DQ dimensions as accuracy, and problems can often either be attributed to multiple sources for the same Data or through judgment being involved. Accessibility forms the second DQ category, including the DQ dimension of the same name and problems being related to the technical accessibility of the Data (Strong et al, 1997). The third DQ category is that of contextual DQ, containing context dependent DQ dimensions such as timeliness, while representational DQ and such dimensions as interpretability form the fourth and final DQ category (Strong et al, 1997).

Data Quality therefore is also dependent on context, leading to its possible definition as a measure of fit with regards to Data usage for business processes (Österle & Otto, 2016). This measure of fitness is also sometimes referred to as suitability (Sautter et al, 2018). DQ is also highly flexible according to Österle & Otto (2016), often deteriorating over time if no actions for its improvement are taken. Insufficient DQ is sometimes only noticed once the Data Quality has deteriorated to a point where it hinders business processes. Therefore, Data Quality and its management have to be considered as a part of overall Data Governance (Österle & Otto, 2016).

1.3 Data Quality Management

Data Quality Management (DQM) consists of the analysis, improvement, and assurance of Data Quality (Österle & Otto, 2016). The first part, analysis, refers to the assessment of the quality of the company's current Data (Österle & Otto, 2016). This Data Quality assessment can be conducted by adopting a variety of Different Data Quality Assessment methodologies, of which thirteen are discussed by Batini et al (2009). The most complete of these according to Batini et al (2009) is constituted from parts of the other methodologies and described as the "Complete Data Quality" Methodology (CDQ). It consists of three phases, the first of which is state reconstruction. This phase consists of the reconstruction of the relationships between the different organizational units, as well as processes, services, and Data (Batini et al, 2009). These relationships are then modelled, in order to visualize the usage of Data and the role it plays in the diverse business processes.

The second phase of CDQ is the assessment phase, which consists of the designation of new Data Quality targets (Batini et al, 2009). This requires the aforementioned knowledge of the usage of Data within the enterprise, since the Data Quality targets have to be set in a way that the Data can support the business processes. Also, since costs and efforts for the improvement of Data are considerable, therefore the Data Quality targets have to focus on the major problems (Batini et al, 2009). The third and final phase of CDQ is the improvement phase (Batini et al, 2009). It includes the identification of root causes of Data Quality issues identified in previous phases. Once these root causes have been identified, they can be addressed, a process for which adequate strategies and methods are chosen. This phase also includes an evaluation of the costs of these improvement measures. (Batini et al, 2009). After the Data Quality has been assessed, a plan can be formulated for its improvement (Österle & Otto, 2016). This process of DQ improvement aims at improving the DQ in the most effective manner, minimizing costs for the necessary improvements (Batini et al, 2009). DQ improvement therefore includes the application of different techniques to the affected Databases, both ones driven by Data and ones driven by processes (Batini et al, 2009). Besides this improvement, DQM also includes the assurance of DQ according to Österle & Otto (2016), referring to the need to improve DQ not only as a reactive measure, but to manage it even before DQ problems make it unavoidable.

Generally, DQM can either be reactive as described by Batini et al (2009) or preventive (Österle & Otto, 2016). Reactive DQM measures focus on the elimination of already existing Data defects, whereas preventive DQM measures focus on the prevention of Data defects before the fact. According to Österle & Otto (2016), reactive DQM is much more common, yet it has a number of disadvantages compared to preventive DQM. For one, there is a risk that the resources necessary for reactive DQM measures might not be available, since the Data defect which triggers their need may appear rather suddenly. Also, reactive DQM often lacks measurement of DQ according to Österle & Otto (2016), leading to a lack of clear DQ targets, which are necessary for effective improvement (Batini et al, 2009). Therefore, there is a differentiation between DQM being something which is triggered by e.g. a DQ problem or as a constant corporate function. This is mimicked in the two DQM frameworks of Total Data Quality Management and Corporate Data Quality Management (Wang, 1998; Österle & Otto, 2016). TDQM assumes a sequential process of definitions, measurement, analysis, and improvement, which is then also implemented in the enterprise. CDQM on the other hand incorporates in akin to a transformation process, which requires consideration on both the strategic, organizational and system layers of an enterprise (Wang, 1998; Österle & Otto, 2016).

1.4 Data Governance

Data Quality Management as a corporate function needs to be embedded into the organizational and decision-making structure of the enterprise, which is referred to as Data Governance (Österle & Otto, 2016). However, Data Governance more generally refers to the need to maximize the value an enterprise gains through its Data, which requires high quality Data but also other considerations, such as (data-)product management. Weber et al (2008, p. 349) therefore describe Data Governance as specifying "the framework for decision rights and accountabilities to encourage desirable behavior in the use of Data", which describes two elements of it. The establishment of accountabilities for DQM includes the implementation of four common roles and the Data Quality board, a board comprised of business unit and IT leaders and deciding on companywide standards and controls (Weber et al, 2008):

(1) **Executive sponsor:** A role referring to the need for top management support for DQM activities. It includes such tasks as providing strategic direction and oversight.

(2) **Chief steward:** A role tasked with implementing the decisions of the Data Quality board, the enforcement of the decided standards as well as the establishment of necessary DQ metrics and goals (Weber et al, 2008).

(3) **Business Data steward:** A Data Governance role referring to the need to adapt the DQ standards or policies to its area of responsibility from a business perspective (Weber et al, 2008).

(4) **Technical Data steward:** The technical counterpart of the previous role, providing e.g. standardized Data elements and formats as well as profiling the details of the source system and the flow of Data between different systems (Weber et al, 2008).

The other element of Data Governance, a framework for decision-making rights, refers to the assignment of responsibilities of specific roles for specific decision-areas such as the planning of Data Quality initiatives (Wende, 2007). Responsibilities are differentiated between being accountable, being responsible, being consulted and being informed, adhering to the general RACI-Governance matrix (see Figure 1).

Roles / Decision Areas	Executive Sponsor	Data Governance Council	Chief Steward	Business Data Steward	Technical Data Steward	...
Plan data quality initiatives	A	R	C	I	I	
Establish a data quality review process	I	A	R	C	C	
Define data producing processes		A	R	C	C	
Define roles and responsibilities	A	R	C	I	I	
Establish policies, procedures and standards for data quality	A	R	R	C	C	
Create a business data dictionary		A	C	C	R	
Define information systems support		I	A	C	R	
...						

R – Responsible; A – Accountable; C – Consulted; I – Informed

Figure 1 Data Governance matrix (Wende, 2007, p.420)

2. Systematic literature review of Data Quality Management references

In order to assess how DQM is implemented by corporations in operational practice, a systematic literature review according to Webster and Watson (2002) was conducted. This review focused on both scientific research contributions as well as practice-based reports. The general process of such a review is subsequently described, as well as the specific search process for DQM in business practice. Concludingly, a short introduction towards general content analysis is given, which was used to analyze the reviewed references.

2.1 Methodology

A structured literature search as proposed by Webster & Watson (2002) is necessary to fully assess a particular field of research. It assures inclusion and consideration of previous research, as well as a structured and extensive research process. For this Webster & Watson (2002) proposed a seven-step process:

(1) **Definition of the research question(s):** In the first step of a structured literature review, the research questions which should be answered by it are defined and described (Webster & Watson, 2002).

(2) **Clarification of key terms and concepts:** The following second step of a structured literature search according to Webster & Watson (2002) is the identification and clarification of key terms and concepts of the research field.

(3) **Identification of relevant journals and conferences:** After the key terms and concepts are clarified, the third step is the identification of Databases, conferences, and authors of relevance for the research question.

(4) **Database search:** In the fourth step, the key terms, and concepts from the second step are utilized in order to search the relevant Databases, conferences, and authors from the third step for relevant literature (Webster & Watson, 2002). The literature identified in this step of the structured literature search is then judged in terms of its relevance regarding the research question, according to at first, its abstract and later its overall content.

(5) **Backwards search:** After it has been deemed suitable for the research, a backwards search is conducted on it in the fifth step, consisting of an assessment of the references utilized by it (Webster & Watson, 2002).

(6) **Forwards search:** The sixed step is then to also conduct a Forwards Search on the relevant literature, assessing the literature which cites the already included research contributions (Webster & Watson, 2002).

(7) **Result visualization:** The seventh and final step of a structured literature analysis according to Webster & Watson (2002) is to visualize the results from the Database-, Backwards- and Forwards-Search, as well as to identify similarities and differences.

Overall, this paper follows these seven steps as postulated by Webster & Watson (2002).

2.2 Literature search

After an introduction to the research topic via Österle & Otto (2016) some preliminary key terms with regards to Data Quality Management were identified (see Table 1). Each of these was applied in both English and German.

Search term 1	**Search term 2**
Data Quality Management	Organization
Information Quality Management	Implementation
DQM	Design
CDQM	Practice
IQM	Strategy
TDQM	System

Table 1 Search terms

These search terms were combined into different search-strings, which were used for the Database search as postulated by Webster & Watson (2002) and ultimately combined into one comprehensive search string (see Table 2).

Search string
((Data Quality Management OR Information Quality management OR Data Quality OR Information Quality OR DQM OR CDQM OR IQM OR TDQM) AND (organization OR design OR implementation OR practice OR strategy OR system))

Table 2 Scientific research Search string

Besides these key search terms, a number of Databases, journals and conferences were found to be possibly relevant. Especially, the Journal of Data and Information Quality (JDIQ) as well as the Data Governance and Information Quality (DGIQ) conference were identified as relevant, as well as the Proceedings of the ACM.

In order to assess DQM in operational practice not only in research contributions, a search was also conducted for practice reports on this issue, utilizing the key search terms from Table 1. In order to account for the informal nature of such corporate reports on DQM, this search was conducted not in literature Databases but in general search engines. In order to account for regional differences as well for variety of results, the three search engines of Google, Bing and Ecosia were selected and searched with several of the determined DQM search keys. The identified references were assessed, based on four exclusion criteria (see Table 3). The fourth criterion however was only applied to research literature, not practice reports. If an identified research contribution or practice report did meet any of these exclusion criteria, it had to be excluded from the literature analysis. After this assessment, the literature was transferred into a literature management program and analyzed.

Exclusion criteria	Description
Criteria 1	The identified literature was not related to the research topic.
Criteria 2	The identified literature was not available in English or German.
Criteria 3	The author was not able to attain a full text copy of the literature.
Criteria 4	The identified literature did not meet research quality requirements: • Use of references • Linked to practical application

Table 3 Exclusion criteria

2.3 Results analysis

The identified contributions towards Data Quality and its management were subjected to qualitative content analysis. According to Elo & Kyngäs (2008), there are two ways to conduct content analysis: inductive and deductive. The inductive approach focuses on combining individual instances into a general statement, while the deductive approach focuses on theory testing (Elo & Kyngäs, 2008). Since the general aim is the creation of an overview over the issue of Data Quality Management in operational practice through structured literature review and content analysis, the inductive approach is adopted. Inductive content analysis consists of three phases:

(1) **Preparation:** This phase starts with the selection of the unit of analysis, followed by making sense of the Data and getting immersed in it (Elo & Kyngäs, 2008).

(2) **Organization:** This phase consists of open coding, category creation and abstraction. Open coding refers to generally making notes while reading the research contribution, while category creation includes the grouping of categories under higher order headings as well as the merging of similar categories (Elo & Kyngäs, 2008). Abstraction refers to the formulation of a general description of the research topic.

(3) **Reporting:** The last phase of reporting refers to the reporting of both the process of analysis and the results (Elo & Kyngäs, 2008).

Therefore, all Data Quality and Data Quality Management research contributions identified in the systematic literature review were subjected to this inductive analysis process.

3. Systematic literature search results

The following chapter features the references identified through the described systematic literature search process. Application of the search string to the Databases yielded a large amount of results, which had to be assessed before the analysis could be conducted. However, over the course of the Database search, the number of identified relevant articles steadily decreased, which according to Webster & Watson (2002) can be seen as sign that the research topic and its aspects have been thoroughly captured during the structured literature review. In total 56 research contributions were identified in this step if the systematic literature review (see Table 4).

Database	Number (55)	Relevant results
Google Scholar	25	Wang et al (1995), Wand & Wang (1996), Weidema & Wesnaes (1996), Strong et al (1997), Grimmer & Hinrichs (2001), Madnick et al (2003), Winter et al (2003), Heinrich & Klier (2006), Shankaranarayanan & Cai (2006), Otto et al (2007), Batini et al (2009), Hüner et al (2009), Madnick et al (2009), Ofner et al (2009), Otto & Hinderer (2009), Weber et al (2009), Lucas (2010, I, Lucas, 2010, II), Schmidt et al, 2010, Otto et al (2012), Glowalla & Sunyaev (2013), Kwon et al (2014), Laranjeiro et al (2015), Schäffer & Beckmann (2018), Houston et al (2018)
EbscoHost	5	Bai (2012), Liaw et al (2014), Bargh et al (2015), Edelen & Ingwersen (2018), Leadbetter et al (2020)
Springer	1	Al-Ruithe, Benkhelifa & Hameed (2019)
AIS-Library	6	Shankaranarayan et al (2003), Otto (2011), Dalmolen et al (2015), Westin & Sein (2015), Schäffer & Stelzer (2017), Zhang et al (2019)
ACM Digital Library	5	Weber et al (2009), Glowalla & Sunyaev (2014), Francisco et al (2017), Shamala et al (2017), Shankaranarayan & Blake (2017)
ResearchGate	4	Cai & Zhu (2015), Jaya et al (2017), Jaya et al (2019),
Scinapse	2	Merino et al (2016), Heinrich et al (2018)
Alexandria	4	Weber et al (2008), Hüner (2011), Falge et al (2012), Frehe et al (2016)
GBV	2	Falge (2014), Österle & Otto (2016)
Swiss National Library	4	Würthele (2003), Weber (2009), Baghi (2017)

Table 4 Database search results

After the Database search, the identified articles were subject to both a backward and a forward search (Webster & Watson, 2002). A backward search refers to the review of the references used by the already selected articles, while a forward search refers to a review of the literature which cites the already selected articles. In total, fifteen new articles were identified over the course of these two steps (see Table 5).

Backwards Search: 10	Forwards Search: 5
Wang & Strong (1996)	Wende (2007)
Wang (1998)	Ofner et al (2013)
Pipino et al (2002)	Kreis (2017)
Ryu et al (2006)	Nurminen (2017)
Batini et al (2007)	Sautter et al (2018)
Caballero et al (2008)	
Otto & Hüner (2009)	
Otto & Ebner (2010)	
Hüner et al (2011)	
Falge et al (2013)	

Table 5 Backward and Forward Search Results

In total 70 articles were included in the literature analysis over the course of the search (see Appendixes D to G). While the number of references excluded due to Criterion 1 was very high, such as articles focusing on technical consideration towards DQ without a link to its management, no articles were excluded due to Criterion 2. Besides this sixteen articles had to be excluded due to Criterion 3, with the author being unable to acquire a full-text copy of the identified research contributions (see Appendix A). Eight were excluded due to Criterion 4. (see Appendix B)

In terms of literature quality, the field seems to be very heterogeneous. On average, the articles in the literature pool were cited 241 times, however in median they were only cited nineteen times, with a standard deviation of 680. The average publication year was 2010, with 2011 in median, with a standard deviation of six. Overall, these figures hint at a fairly recent literature pool, which however is very heterogenous in terms of importance to the research field (see Table 6). Some articles seem to form important centerpieces of the research field with a high amount of citations, while the majority of articles seem to be much less important for the research field and only cited sparsely.

Year	**1995**	**1996**	**1997**	**1998**	**2002**	**2003**	**2006**	**2007**	**2008**	**2009**	
Sources (∑)	1	3	1	1	1	4	3	3	2	9	
Citations (∑)	901	7414	1652	1208	1919	303	293	226	85	1878	
Year	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Sources (∑)	4	5	3	3	4	4	3	8	5	3	1
Citations (∑)	55	164	107	34	517	105	148	71	46	17	2

Table 6 Search Results

Also, the Basket of Eight is also almost not presented in the literature pool, with the most frequent journal being the Journal of Data and Information Quality which is not ranked in the VHB index. The only article ranked in the VHB is Madnick et al (2003), which was published in the Journal of Management Information System, an A ranked journal. On the level of B ranked journals the journal of Business & Information Systems Engineering is also represented via Zhang et al (2019) and Glowalla & Sunyaev (2013).

The literature pool also contains a number of dissertations and other grey literature, however none of these seems to be in immediate and serious conflict with the research quality criteria according to Balzert et al. (2017). As described before, a search for practical reports regarding DQM in corporate practice was also conducted in three search Engines, Google, Bing and Ecosia. Besides the recommended practice report of Trumpetter (2015) regarding DQM at the Stadtwerke Munich, seven additional practice reports where identified in this way (see Table 5).

Practice reports
Trumpetter Joerg. (2015). Etablierung eines nachhaltigen Datenqualitätsmanagements bei den Stadtwerken München.
Everding (2010): Pfalzwerke Netzgesellschaft mbH | SAP Daten analysieren, Datenqualität optimieren – intuitive Analysen ohne Suchvorgaben (https://www.infozoom.com/sap-daten-datenqualitaet-pfalzwerke-netzgesellschaft/)
ACT IT-Consulting & Services GmbH (2011): Datenqualitätsmanagement bei der Barmenia Versicherung (https://www.pressebox.de/pressemitteilung/act-gruppe/Datenqualitaetsmanagement-bei-der-Barmenia-Versicherung/boxid/415759)
Meyer (2012): SWB Energie und Wasser | Datenqualitätsmanagement SAP Daten: Datenkontrolle im Fachbereich (https://www.infozoom.com/datenqualitaetsmanagement-sap-daten-swb-energie-und-wasser/)
Meyer (2015): Stadtwerke Winsen (Luhe) GmbH | Datenqualitätsmanagement: Wertschöpfung durch saubere Daten (https://www.infozoom.com/datenqualitaetsmanagement-stadtwerke-winsen-luhe-gmbh/)
Eggheads GmbH (2020): Hartje: Jetzt mit zentralisierten Produktdaten Vollgas geben (https://www.pressebox.de/pressemitteilung/eggheads-gmbh/Jetzt-mit-zentralisierten-Produktdaten-Vollgas-geben/boxid/1001973)
Finanz Informatik (2020, I): Datenqualität ist messbar (https://www.f-i.de/News/ITmagazin/Archiv/2018/Integrierter-Datenhaushalt/Titelthema/Datenqualitaet-ist-messbar)
Finanz Informatik (2020, II): Neues vom Integrierten Datenhaushalt (https://www.f-i.de/News/ITmagazin/Archiv/2019/FI-Connect-2019-VorausDenkenMachen/Loesungen-Praxis/Neues-vom-Integrierten-Datenhaushalt)

Table 7 Practice reports

However, compared to the research contributions only a small number of practice reports was found, with the recommended one being the most detailed. The nature of the recommended report as well as the experiences during the search for practice reports hints a lack of accessibility in this regard. The practice report of Trumpetter (2015) was made at a practical lecture at the University of Passau and could not be found in generic or literature databases. Adding to this, the term practice report was often found in the context of brochures for industry conferences or lecture's during this part of the search, however none of these published any of the mentioned reports. The majority, most detailed and freely available practice reports were therefore found in scientific research contributions towards DQM, such as by Österle & Otto (2016), Weber (2009) or Falge (2014).

4. Analysis of Data Quality Management research contributions

After the search for and selection of suitable references, an analysis and assessment of the research contributions is conducted, with a focus on the utilized methods and theoretical foundations. The 70 references identified over the course of the literature search were analyzed in order to identify both individual and overarching concepts with regards to Data Quality and Data Quality Management (see Appendixes D to G). Besides that, they were also analyzed in terms of their goals, theoretical foundations and utilized research methods

4.1 Analysis elements

4.1.1 Theoretical foundations

With a research history of around 25 years, Data Quality Management has accumulated a considerable number of theories related to it and its research. Data Quality Management research therefore is usually either built upon an existing theoretical foundation or tries to establish such a theoretical foundation. Over the course of the literature review, the mayor theoretical foundation of the Resource Based View was identified. Besides this however, many research contributions did not explicitly refer to a clear and well-developed theory as their foundation.

5.1.1.1 Resource Based View

A major theoretical foundation for the research contributions regarding DQM is the Resource Based View. Kwon et al (2014) describe RBV as differentiating between tangible and intangible resources. Elements such as the IT system are seen as tangible resources, while elements such as skills and knowledge, e.g. with regards to DQM, are seen as intangible resources. Competitive advantages are derived from an increased IT Capability through both tangible and intangible resources. Baghi (2017) similarly bases its research into **Data Quality** controlling on the RBV, outlining both organizational capabilities and routines and the role they play in achieving a competitive advantage.

5.1.1.2 Foundations not clear

Besides these described DQM research foundation, most of the identified research contributions did not refer to a clear theoretical foundation as the basis of their DQM research, such as Wang & Strong (1996). They refer to management concepts such as Total Quality Management or to standards such as ISO9000, but do not explicitly to a clearly defined theory such as the Resource-Based-View.

4.1.2 Research methodologies

Besides the theoretical foundation for Data Quality Management research, the utilized methodologies are also of great importance for a comprehensive overview over this research topic. Both Madnick et al (2009) and Jaya et al (2019) describe the different research methodologies prevalent in Data Quality research. During the structured literature review, the four main research methodologies were those of case study, action research, design science and survey, often combined with each other and minor methodologies in the reviewed research contributions. Besides these, a number of minor, less developed research approaches such as systematic literature review were also identified, as well as a number of references where the utilized methodology or research approach was unclear.

4.1.2.1 Case study

Case study research is one of the most prominent research methodologies in DQM research besides design science. The research target is observed within its natural setting, without any experimental variable controls (Benbasat et al, 1987). The methodology is also described by Dresch et al (2015) as a six-step research process:

(1) Definition of a Theoretical Conceptual Structure: This first step of case study research consists of a mapping of the relevant literature as well as the identification of research gaps. According to Dresch et al (2015) this also includes the extraction of the constructs that should be verified by the research, concluding in the definition of the objectives and boundaries of the research project.

(2) Case Planning: After the structure for the case has been defined, the overall case is planned. The target and means of Data collection and Data analysis are chosen, as well as a protocol for Data collection developed (Dresch et al, 2015).

(3) Pilot Test Driving: Before the actual Data collection a pilot test is conducted by the researcher. Application procedures are verified, besides identifying the association between the obtained Data and the research constructs (Dresch et al, 2015).

(4) Data Collection: The Data collection step is started by contacting the main informants as well as by providing an estimate for the time necessary for the case (Dresch et al, 2015). The main means of Data collection in case studies are interviews, therefore the researcher effect has to be accounted for and limited.

(5) Data Analysis: After the Data has been collected, it is analyzed, identifying causalities (Dresch et al, 2015). It also consists of the building of a case narrative, as well as of Data reduction.

(6) Report Creation: Case study research concludes in the submission of an overall case report. This report should refer to the theoretical foundation but not adjust the foundation to the results (Dresch et al, 2015).

The main disadvantage of case study is the impossibility to control the variables, but it also does not require the active participation of the researcher (Dresch et al, 2015). However, while the general goal of case study research is the generation of new theories and factor relationships, some of the reviewed research contributions utilize their cases more akin to detailed and specific examples.

4.1.2.4 Action research

Dresch et al (2015) also describe the methodology of action research, which consist of six overall steps. However, in difference to case study the process follows a circular approach rather than a sequential one, with the researcher being a part of the partner organization during the research. The main action research phase is preceded by a preliminary round to define context and purpose, which leads over to the first iteration of the action research cycle. According to Coughlan & Coglan, (2002) these steps of action research include:

(1) Data collection: The first step of the action research cycle is the collection of the necessary Data, both hard, statistical Data and soft, perceptual Data (Coughlan & Coglan, 2002).

(2) Data Feedback: After the Data has been collected, feedback from the partner organization is gathered (Coughlan & Coglan, 2002).

(3) Data Analysis: The collected Data is analyzed in collaboration of the researcher and the partner organization, with the analysis criteria and tools being related to the research goal (Coughlan & Coglan, 2002).

(4) Action planning: After Data has been collected and analyzed, actions are planned, based on the insight generated from the Data. Like the Data analysis, this step is done in collaboration, defining what has to be done, by whom and when (Coughlan & Coglan, 2002).

(5) Action implementation: The formulated action plan is then implemented at the partner organization in the manner specified in the plan. This is still done in collaboration of the researcher and the partner organization (Coughlan & Coglan, 2002).

(6) Action evaluating: The last step of the action research cycle, consisting of an evaluation of the action results. Both positive and negative results are evaluated in this step (Coughlan & Coglan, 2002).

If necessary, the circle is then restarted with another Data collection. Besides that, is also a (meta)-phase of monitoring the six steps. Action research can be conducted without predefined constructs, which is not possible in case study according to Dresch et al (2015).

4.1.2.3 Design science

The most prevalent research methodology identified in the reviewed research contributions is the Design science methodology. In difference to case study and action research, design science puts a heavier focus not only on understanding the problem but on providing practical solutions (Dresch et al, 2015). Therefore, artifact creation is one of the main goals of design science, as its theory formulation according to Dresch et al (2015). Due to this, it may be applied in case of a lack of a clearly developed theoretical foundation. Also, Design science is dependent on six elements which are described by both Dresch et al (2015) and March & Storey (2008) as:

(1) Problem: Problem as a design science element refers to the need for the problem to be relevant, as well as it being clearly described (March & Storey, 2008; Dresch et al, 2015).

(2) Solution: The solution element of design science refers to the need that the researcher presents the lack of an artifact that addresses the problem as well as presents that the designed artifact does so (Dresch et al, 2015).

(3) Development: An element of design research that refers to the need for artefact/solution to be properly developed (Dresch et al, 2015).

(4) Evaluation: After the artifact has been developed, it has to be evaluated if the artifact meets the requirements of utility and viability, which is what this element refers to (Dresch et al, 2015).

(5) **Value addition:** A design science element that refers to the need of the developed artifact to add new knowledge and provide a useful solution (Dresch et al, 2015).

(6) **Communication:** The last design science research element refers to the need for the researcher to communicate both the results of the research and how it was achieved (Dresch et al, 2015).

For this design science-based theory formulation four phases are postulated, with (1) solution incubation being the first. This is followed by (2) solution refinement, (3) the formulation of mid-range theories and (4) the formulation of formal theories (Dresch et al, 2015). While mid-range theories are context dependent, formal theories may be applied regardless of the context.

4.1.2.4 Survey

Besides the three research methodologies already presented, a number of identified contributions also utilized some kind of survey as one of its research methodologies (see Appendix C). Both Wang & Strong (1996) and Strong et al (1997) conducted surveys as a part of their research, as did others such as Otto & Ebner (2010), Kwon et al (2014), Kreis (2017) and Shamala et al (2017). According to Gable (1994), surveys as method put an emphasis on the quantitative approach, focusing on the collection of Data from a large number of organizations. Its approach emphasizes objectivity and testability, but also limits the degree of understanding of the subject which can be achieved. Gable (1994) also describes surveys as being preferable in subject areas which are less accessible to methods of field work, such as case studies. With regards to conducting both qualitative and statistical surveys as a research method, Jansen (2010) formulates four steps:

(1) **Definition of the knowledge goals:** In this first step of survey research, both the topic and its researched aspects are defined, as well as its empirical domain. Besides these, the unit of Data collection and the knowledge function are also defined, with this step being identical for qualitative and statistical surveys (Jansen, 2010).

(2) **Sampling:** After the knowledge goals are defined, sampling is conducted in the second step, with the method of selection and criterion for size being defined. Qualitative surveys select for diversity and with the coverage of the population size being the criterion for size. Statistical surveys on the other hand select for probability, with the criterion for size being the estimate precision (Jansen, 2010).

(3) Data collection: In the third step, Data is collected, by questioning people but also by the observation of interactions or artifacts (Jansen, 2010).

(4) Analysis: In the last step of survey research, the collected Data is analyzed on three different layers.

  a. First level analysis focuses on a unidimensional description, by coding in qualitative surveys and frequency counting and descriptive statistics in statistical surveys (Jansen, 2010).

  b. On the second level, the Data is described multidimensional, which is done case-oriented and/or concept-oriented in qualitative surveys and unit-oriented and/or variable oriented in statistical surveys.

  c. On the third and final level of explanation, qualitative surveys analyze relationships between different types and contextual conditions, while statistical surveys explain the gradual variation in the dependent variable (Jansen, 2010).

4.1.2.5 Minor methods

Besides these four main research methods in DQM research contributions, several contributions make use of minor methods of Data collection such as systematic literature reviews. These are less well developed and founded than the four described main methods found in DQM research contributions.

4.1.3 Research quality criteria

Another issue that has to be considered for an overview of Data Quality Management is the quality of its research contributions. Research quality can be judged among other things by adherence to quality criteria as the twelve established by Balzert et al (2015):

(1) Honesty: A research quality criterion that can be seen as the absence of e.g. deception or plagiarism (Balzert et al, 2015). Due to this, it can be seen as an absolute criterion, with no acceptable degree of a lack of honesty.

(2) Objectivity: Objective research is conducted from a neutral stance, eliminating such factors as the researcher's predispositions with regards to the research results (Balzert et al, 2015).

(3) **Verifiability:** This research quality criterion refers to the need for the research results being possible reproduced in principle by other researchers. Therefore, all relevant factors have to be laid out and no Information hidden or unmentioned (Balzert et al, 2015).

(4) **Reliability:** A more mathematical research quality criterion, referring to the degree that the utilized methods measure precisely and that the results are stable. Stable results mean that other researchers utilizing the same methods and tools should get the same results.
(5) **Validity:** The research quality criterion of validity refers to the utilized method measuring what it is supposed to, rather than how precise it measures as in reliability.
(6) **Understandability:** A more subjective research criterion, referring to the usage of standardized research contribution parts, such as the use of a table of contents and appendixes.
(7) **Relevance:** Relevant research contributions provides new insight towards the research field, contains high value Information, and helps in the solution of practical problems.
(8) **Logical argumentation:** This refers to the usage of logical arguments, being based upon premises. A logical argument finishes with a conclusion and can either follow inductive or deductive reasoning.
(9) **Originality:** A research quality criterion referring to the need for research contributions to be provided by the researcher and being new.
(10) **Traceability:** This research quality criterion is dependent on the fulfilment of the nine previous criteria and refers to the reader of the research contribution being able to trace the process of result generation.
(11) **Fairness:** A more recent research quality criterion which refers to the researcher's behavior as a part of the scientific community.
(12) **Responsibility:** The last research quality criterion, which refers to responsibility towards the researcher herself and her team, but also towards the scientific community.

However, only a number of these criteria can be seen as relevant in the context of evaluating the quality of the reviewed literature. Fairness and Responsibility cannot be judged within the context of this paper, requiring further insight into the context of the contributions or due to considerable time having been passed since the publication. This also applies to the Honesty criterion and the Objectivity criterion.

Therefore, the eight criteria of (1) Verifiability, (2) Reliability, (3) Validity, (4) Understandability, (5) Relevance, (6) Logical argumentation, (7) Originality and (8) Traceability were utilized in order to assess the research quality of the reviewed contributions.

4.2 Analysis and assessment of DQ research contributions

Many of the identified contributions referred to specific Data Quality concepts which are also relevant with regards to its management, such as the effects of Data Quality, Data Quality dimensions and Data Quality improvement. These scientific research contributions are analyzed in terms of (1) their goals, (2) the utilized research method, (3) their theoretical foundation and (5) results, as well as being assessed. Due to the large number of research contributions, only the most important research contributions towards each topic will be analyzed and discussed in detail, with the rest being summarized. Importance is judged upon (a) relevance for the overall topic of Data Quality Management in practice and (b) research importance based on citations.

4.2.1 Research contributions related to DQ effects

One concept identified over the course of the structured literature review was that of the effects of Data Quality. Four sources from the identified literature concerned themselves with this issue.

Research contribution	**Research Goal**	**Foundation**	**Method**
Falge et al (2012)	**DQ requirements for BPs**	**Business Processes & Data Quality**	**Qualitative content analysis**
Kwon et al (2014)	**Influence factors on Big-Data analytics adoption**	**Resource-Based-View & Isomorphism**	**Survey**
Dalmolen et al (2015)	Product Information sharing	No clear foundation is provided	Case study
Kreis (2017)	Success factors for Data Migration	No clear foundation is provided	Survey

Table 8 Research contributions related to Data Quality effects

Exemplarily, the research contributions of Falge et al (2012) and Kwon et al (2014) are analyzed.

Falge et al (2012). Data Quality Requirements of Collaborative Business Processes

Falge et al (2012) aim at identifying the requirements that collaborative business processes pose for DQ. They describe the background of their research in the terms of business processes, collaborative business processes and Data Quality. However, no explicit theoretical foundation is mentioned as the basis of their research contribution. In order to achieve their goal Falge et al (2012) conducted a qualitative content analysis of twelve case studies. A differentiation was made between Supply Chain, Commerce, Maintenance & Repair and Finance business processes. Regarding Supply Chain business processes, it was found that the Data Quality dimensions of timeliness, accuracy and completeness were most important, which coincided with the results for Commerce Business Processes, only differing in which Data classes were most important with regards to Data Quality. The results for Maintenance and Repair processes were similar, with an additional important Data Quality dimension being Temporal Validity (Falge et al, 2012). Finance business processes, however, were found to be more dependent on the Data Quality dimensions of accuracy, Data security as well as the usual consistency.

Assessment: The research contribution provided by Falge et al (2012) provides important insight with regards to which DQ dimensions are seen as important for different business processes by practitioners. It achieves this via a well-documented research process of qualitative content analysis, which however in itself is only a minor research method. Therefore, the only significant flaw from a theoretical perspective is the lack of a well-established theoretical foundation for Falge et al (2012).

Kwon et al (2014). Data Quality Management, Data usage experience and acquisition intention of big Data analytics.

Kwon et al (2014) aimed at explaining the factors which influence a company's intention to acquire big Data analytics tools. The research contribution is based upon the resource-based view as a theory and as well as on the concept of isomorphism. In order to validate the proposed model of adoption influence factors, a survey was conducted. This survey was conducted on procurement specialists and included several minor adjustments. In total 306 responses were collected, yielding a total of 18 survey measures for the six model constructs. The model proposed by Kwon et al (2014) with regards to acquisition intention of Big Data analytics postulates that this intention is influenced by the resource facilitating condition as well as the perceived benefits from the usage of both internal and external Data. The perceived benefits were then dependent on the degrees of Data consistency and completeness. They found that Data Quality, expressed as Data consistency and completeness, had an overall positive impact on the decision to adopt Big Data analytics (Kwon et al, 2014). Data consistency and completeness both had a positive influence on the perceived benefits from internal and external Data usage. Interestingly, only perceived benefits from external Data usage has a positive effect on the acquisition intention regarding big Data analytics, with perceived benefits from internal Data usage having a negative link to acquisition intention (Kwon et al, 2014).

Assessment: The research contribution of Kwon et al (2014) provides an interesting insight into how Data Quality dimensions influence the adoption of Big Data analytics. Their research is well founded theoretically in the resource-based view and the concept of isomorphism. Their results are tested both in terms of validity and reliability, with no other apparent issues with regards to research quality.

Other DQ effect references

Dalmolen et al (2015) investigated product Information sharing between 22 Dutch companies and found that Data Quality is of vital importance in order to enable collaboration and Information sharing between these companies. Another view on the effects of Data Quality is presented by Kreis (2017), who investigated the influence of Data Quality on Data migration projects. It was found that poor Data Quality hinders Data migration, with good Data Quality as well as Data governance being key influencing factors for the success of Data migration.

4.2.2 Research contributions related to DQ assessment

Yet another important concept with regards to Data Quality Management identified in the review was the assessment of Data Quality, which as previously described is an important part of Data Quality Management (Österle & Otto, 2016). In total, ten references identified in the literature review concerned themselves with Data Quality assessment in one way or another.

Research contribution	Goal	Foundation	Method
Weidema & Wesnaes (1996)	DQ assessment in Life-Cycle Inventories	No clear foundation is stated	No clear method is given
Pipino et al (2002)	Combination of subjective and objective DQ assessment approaches	No clear foundation is stated	No clear method is given
Batini et al (2007)	Definition of an DQ assessment methodology	No clear foundation is stated	No clear method is given
Batini et al (2009)	Description and comparison of DQ assessment methodologies	No clear foundation is stated	Comparative analysis

Otto & Ebner (2010)	Survey of DQ assessment in companies	Data as a Product	Survey
Cai & Zhu (2015)	Overview over DQ and DQ assessment challenges through Big Data	No clear foundation is stated	Literature review
Laranjeiro et al (2015)	Overview over State of the Art in the classification of poor Data and mapping of Data Quality problems	No clear foundation is stated	Literature review
Merino et al (2016)	Provision of model for DQ assessment of Big Data	No clear foundation is stated	Design Science
Edelen & Ingwersen (2018)	Revision to the assessment of DQ in life cycles	No clear foundation is stated	No clear method is given
Zhang et al (2019)	Creation of an approach for DQ assessment	Semiotic theory	Design Science

Table 9 Research contributions related to Data Quality assessment

Exemplarily, the research contributions of Pipino et al (2002) and Batini et al (2009) are analyzed.

Pipino, L. L., Lee, Y. W., & Wang, R. Y (2002). Data Quality assessment

The research of Pipino et al (2002) aims at presenting an approach which merges subjective and objective assessments of Data Quality. For this however, neither a clear and explicit research foundation nor an explicit method are stated. The development of the approach shares similarities with Design Science in its development but lacks any formal statements and process steps. The presented approach is illustrated in two case studies. Pipino et al (2002) establish categories of metrics for the assessment of these dimensions and conclude in establishing three steps necessary to improve Data Quality (Pipino et al, 2002). The first of these is the conducting of both subjective and objective Data Quality assessments; those results can fall in any of four Data Quality results quadrants. This is followed in the second step by result comparison, discrepancy identification and root cause analysis (Pipino et al, 2002). The last step according to Pipino et al (2002) then forms the determination and planning of corrective actions.

Assessment: Overall Pipino et al (2002) present an unorganized research contribution with both unclear theoretical foundations and methods. Its approach can be described as merely descriptive. Ultimately the lack of clear foundation and well as of the application of a clear method puts the research contribution in doubt.

Batini, C., Barone, D., Mastrella, M., Maurino, A., & Ruffini, C. (2009). Methodologies for Data Quality assessment and improvement

Batini et al (2009) aim at providing an overview over methodologies for Data Quality assessment as well as its improvement. For this no explicit research foundation is provided by Batini et al (2009). For the description and differentiation between different methodologies for DQ assessment, comparative analysis is used. In total, thirteen methodologies are described and explained. They differ with regards to the Information system to which they should be applied to as well as to for which types of Data they are most suitable. Data is differentiated between structured and unstructured Data, whereas Information systems are divided into distributed Information systems, Data Warehouses, Cooperative Information systems and the Web (Batini et al, 2009).

Assessment: Overall, the article of Batini et al (2009) provides important overviews over different methodologies which can be utilized in the context of Data Quality Management in order to assess and improve Data Quality. It suffers however from the lack of a theoretical foundation for research, as well as through its limitation of comparative description of the different methodologies.

Other DQ assessment research contributions

The topic of Data Quality assessment with regards to life cycle Data is discussed by Weidema & Wesnaes (1996), with a strong consideration of Data Quality factors and Data Quality indicators. These indicators are reliability and completeness, as well as temporal, geographical and technological correlation. However, their combination with five scores also hints at a utilization more akin to Data Quality metrics. Ultimately, Weidema & Wesnaes (1996) combine these indicators with uncertainty, further hinting at an understanding more akin to Data Quality metrics. Otto & Ebner (2010) describe the mentioned Data Quality dimensions of accuracy, timeliness, and completeness, as well as consistency, relevancy, and accessibility (Otto & Ebner, 2010). Consistency refers to the degree that data in one database correspond to the same data in another database. Relevancy refers to the data being usable for the intended purpose, while accessibility refers to the ability to access that data at any given point in time (Otto & Ebner, 2010).Batini et al (2007) formulate a methodology for Data Quality assessment, consisting of the four steps of (1) DQ risk prioritization, (2) DQ risk identification, (3) DQ risk measurement and (4) DQ risk monitoring (Batini et al, 2007). Besides these, Data Quality Assessment is also discussed by Laranjeiro et al (2015) in the context of the identification of “poor” Data. They concluded in a map of common Data Quality problems, differentiating between single-source and multi-source problems.

A more systematic and research-oriented approach towards Data Quality assessment is adopted by Cai & Zhu (2015). They develop a process for dynamic Data Quality assessment with regards to big Data. This dynamic Data Quality assessment process starts with the determination of the goals of Data collection, followed by the determination of relevant quality dimensions and elements (Cai & Zhu, 2015). From this determination the Data Quality indicators can be determined, which are in turn used for the formulation of a baseline for the Data Quality evaluation. After the Data Quality assessment, the Data can be compared with the established baseline for it. Should the Data be found unsatisfactory compared to this baseline, new Data has to be collected. Otherwise, the Data Quality assessment process described by Cai & Zhu (2015) continues into the Data analysis phase. This Data analysis phase, while not technically part of the Data Quality assessment, is nevertheless important with regards to the dynamic nature and necessary adjustment of the Data Quality assessment process (Cai & Zhu, 2015).

Similarly, Merino et al (2016) adopted a model for the assessment of Data Quality in use for big Data projects. Their 3-A's model is based on three Data Quality characteristics or three A's. These three characteristics are (1) contextual adequacy, (2) temporal adequacy, and (3) operational adequacy. Besides that, Data Quality assessment was also addressed by Edelen & Ingwersen (2018), who dealt with life-cycle Data and the assessment of its quality. For that, they differentiate between flow and process indicators, ultimately updating the usually used pedigree matrix for LCA (Edelen & Ingwersen (2018). More recently, Data Quality Assessment has also been discussed by Zhang et al (2019). While most approaches towards Data Quality assessment focus on explored and owned Data, their approach focuses on the assessment of the quality of free, unowned Data sets. This approach is referred to as LANG and consists of the two stages of semantic and syntactic, as well as several checks (Zhang et al, 2019).

4.2.3 Research contributions related to DQ metrics

Tying into the topic of Data Quality assessment is that of Data Quality metrics, which enable the measurement of Data Quality.

Research contribution	**Goal**	**Foundation**	**Method**
Würthele (2003)	Development of DQ metrics	No explicit theoretical foundation is given	Design Science
Hüner (2011, I)	Provision of a method for DQ metric specification	Business Engineering	Design Science
Hüner et al (2011)	Analysis of Data use at the case company and proposition of DQ metrics	No explicit theoretical foundation is given	Case study
Bai (2012)	**Development of a mathematical framework for DQ impact and DQ metrics**	**No explicit theoretical foundation is given**	**Case study**
Heinrich et al (2018)	**Provision of requirements for DQ metrics**	**No explicit theoretical foundation is given**	**Unclear**

Table 10 Research contributions related to Data Quality metrics

Exemplarily, the research contributions of Bai (2012) and Heinrich et al (2018) are analyzed.

Bai, X. (2012). A Mathematical Framework for Data Quality Management in Enterprise Systems

The goal of the research contribution of Bai (2012) is the provision of a mathematical framework for Data Quality impact and Data Quality metrics. No explicit research foundation is provided. The research method described as mathematical modeling, but the framework is also evaluated on a case study. Generally, Bai (2012) presents a framework for modeling, part of which are Data Quality metrics regarding the impact of Data Quality errors. These error metrics include the error incidence rate, referring to the total number of error incidences, and the proportion of net monetary error, referring to the monetary magnitude due to the discrepancy between the actual Data value and the recorded value (Bai, 2012).

Assessment: Overall Bai (2012) provides an interesting research contribution towards DQ metrics. It is however limited due to the lack of a clear foundation for its research.

Heinrich, B., Diana Hristova, Mathias Klier, Alexander Schiller, & Michael Szubartowicz (2018). Requirements for Data Quality Metrics

The goal of Heinrich et al (2018) is the provision of requirements of Data Quality metrics, however no explicit research foundation is provided. The utilized methods remains also unclear. Five specific requirements for Data Quality metrics were postulated (Heinrich et al, 2018). The first is the existence of minimum and maximum metric values, while the second one requests that metric values have to be interval-scaled. According to the third definite requirement for Data Quality metrics, determination of the configuration parameters of the quality metrics must be possible, based upon the three criteria of objectivity, reliability, and validity. The fourth requirement is that of a sound aggregation of the metric values, while the fifth and final requirement being the economic efficiency of the metric (Heinrich et al, 2018).

Assessment: While the requirements were found to be practical and possible to fulfill by Heinrich et al (2018), they are limited through the lack of both an explicit theoretical foundation and a clear method by which they were developed.

Other DQ metrics research contributions

Hüner (2011) developed a method for the specification of business-oriented Data Quality metrics. According to this method, it is the task of a Data Quality metric to monitor Data Quality defects before they become an issue for the business operations. For Data Quality, metric specification knowledge of the requirements for the Data has be acquired. These requirements also have to be made measurable (Hüner, 2011; I). The method consists of three overall phases, those being (1) Information collection, (2) Analysis & Specification as well as (3) Approval and Documentation. Besides these, Hüner (2011, I) also distinguishes between six roles with different responsibilities with regards to Data Quality metric specification (Hüner, 2011). These six roles are the Corporate Data Steward, the Sponsor, the Process owner, the Data user, the Business Data Steward, and the Technical Data Steward. Ultimately, the application of the proposed method for Data Quality metric specification enables an objective assessment of Data Quality, although according to Hüner (2011) the considerable effort for its application should be considered as well.

Würthele (2003) provides an extensive contribution towards this issue by first developing a Data Quality radar, encompassing all aspects of Data Quality. A process model is presented as well, being based on an expanded notation for Data. This is then further developed into a system of Data Quality metrics through aggregation. Tying into this, Hüner et al (2011) provided yet another contribution to this research area, with a focus on the quality of product Data in supply chains. Besides identifying Data defects, their research also focused on the specification and application of Data Quality metrics to their respective case study at Beiersdorf. This specification was the subject of the fourth and final phase of Beiersdorf's Data Quality project, preceded by the identification and analysis of critical Data defects (Hüner et al, 2011). The measurement interval for the identified Data Quality metrics was then normalized to be either one, for no Data object containing a critical Data defect, and zero, for all Data objects containing a critical defect. The overall seven Data Quality metrics, such as Bill of Materials, and their values were to be calculated monthly, based on a total of 32 validations rules (Hüner et al, 2011).

4.2.4 Research contributions related to DQ dimensions

Another topic prevalent in the literature and already hinted at with regards to Data Quality assessment and Data Quality metrics is that of the Data Quality dimensions.

Research contribution	**Goal**	**Foundation**	**Method**
Wand & Wang (1996)	Identification of DQ dimensions	Information System	Unclear
Wang & Strong (1996)	**Develop a framework for capturing DQ aspects**	**Data as a Product**	**Survey**
Strong et al (1997)	Generating insight regarding how Data Quality is perceived	No explicit foundation is given	Case study
Shamala et al (2017)	**Identify relevant DQ dimensions for Information Security Risk Management**	**Information Quality**	**Comparative Analysis & Survey**

Table 11 Research contributions related to Data Quality dimensions

Exemplarily, the research contributions of Wang & Strong (1996) and Shamala et al (2017) are analyzed.

Wang & Strong (1996). Beyond Accuracy: What Data Quality Means to Data Consumers

The goal is the development of a framework for capturing DQ aspects. The research of Wang & Strong (1996) is based upon the concept of Data as a Product, but with no explicit theory as its foundation. It utilizes a two-stage survey as its main method, with the first stage producing a number of Data Quality attributes. The first stage was conducted on 25 Data consumers as well as 112 MBA students and identified 179 DQ attributes. The second stage assessed the importance of these attributes to Data consumers and was conducted on 1500 MBA students from which 355 viable answers were received. Through factor analysis, a total of twenty Data Quality dimensions were derived from these. Via a two-phase sorting study, these dimensions were sorted into four categories. Wang & Strong (1996) develop a framework for the organization of Data Quality dimensions. This framework consists of four categories of Data Quality, each referring to different Data Quality dimensions. The first of these categories is intrinsic Data Quality, referring to such dimensions as accuracy, followed by contextual Data Quality, referring to such dimensions as timeliness. Representational Data Quality, referring to such dimensions as interpretability, forms the third category, followed by the fourth category of accessibility Data Quality, referring such DQ dimensions as access security (Wang & Strong, 1996).

Assessment: Wang & Strong (1996) provide an important and early contribution towards the organization of Data Quality dimensions. While lacking in clear theoretical foundation it nonetheless provide an important research contribution through their rigorous research process.

Shamala, P., Ahmad, R., Zolait, A., & Sedek, M. (2017). Integrating Information Quality dimensions into Information security risk management (ISRM)

This research contribution aims to identify relevant DQ dimensions for Information Security Risk Management. It however lacks an explicit theoretical foundation. In order to achieve its goals Shamala et al (2017) utilize both comparative analysis and a survey. They utilize a total of thirteen Data Quality dimensions from the comparative analysis for their research model. The survey was used in order to validate the research model, with a total of 150 responses having been received. These responses there subjected to both convergent and discriminant validity analysis was conducted. After this, the model was analyzed via structural model evaluation. Their research concluded in six Data Quality dimensions which significantly influenced the quality of Information gathering for ISRM. These six dimensions were accuracy, amount of Data, completeness, objectivity, reliability, and verifiability (Shamala et al, 2017).

Assessment: The research contribution of Shamala et al (2017), while lacking in theoretical foundation, nonetheless employs a rigorous research process. Its findings provide important insights into a special area of Data Quality, one not often considered with regards to Data Quality Management.

Other DQ dimension research contributions

Wand & Wang (1996) formulate a Data Quality model and its fundamental principles. The Data Quality dimensions are based on ontological functions, but also three design deficiencies are identified with regards to Data Quality dimensions in the Information system context. These are (1) incomplete representation, (2) ambiguous representation and (3) meaningless states (Wand & Wang, 1996). Furthermore, a differentiation is made between intrinsic and extrinsic Data Quality dimensions, commonly found in Data Quality literature. Strong et al (1997) on the other hand revisit the issue of Data Quality dimension categories, differentiating between the different categories of (1) intrinsic, (2) accessibility, (3) contextual and (4) representational DQ.

4.3 Analysis and assessment of DQM research contributions

4.3.1 Research contributions related to DQM and Data Governance

Another concept present within the reviewed literature was that Data Quality Management and Data Governance, with eight identified references containing this concept (see Table 12).

Research contribution	**Goal**	**Foundation**	**Method**
Wende (2007)	Proposal of a model documenting the company-specific decision-making framework of DQM	No clear foundation is given	No clear method is described
Weber et al (2008)	Identification of a Data governance structure with the emphasis on collaboration between business and IT	No clear foundation is given	Case study
Weber (2009)	Development of a reference model for Data Governance	No clear foundation is given	Design Science
Weber et al (2009, I)	Showing possibilities for the organization of DQM and answering open questions	No clear foundation is given	No clear method is described

Weber et al (2009, II)	**Starting a scientific Assessment on Data governance via the transfer of IT governance and organizational theory concepts to Data Governance**	**No clear foundation is given**	**Action research**
Otto (2011)	**Reporting on results of a case study of Data Governance in two large telecommunications companies**	**No clear foundation is given**	**Case study**
Liaw et al (2014)	Expansion and update of a systematic review of clinical governance	No clear foundation is given	Systematic literature review
Al-Ruithe et al (2019)	Provision of an overview over Data and Cloud Governance articles	No clear foundation is given	Systematic literature review

Table 12 Research contributions related to DQM and Data Governance

Exemplarily, the research contributions of Weber et al (2009, II) and Otto (2011) are analyzed.

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All---A Contingency Approach to Data Governance

Weber et al (2009, II) aim at transferring concepts of IT governance and organizational theory to Data Governance. They lack an explicit theoretical foundation, only describing the background of the research contribution in terms of the concepts of DQM, IT Governance and Data Governance. Their research is conducted in a community action research project, with the topic of the contribution being the subject of a dedicated workshop. Weber et al (2009, II) develop a Data governance model as well as demonstrating the influence of other factors. Their Data Governance contains the DQM decision areas as well as main activities, assigning roles to them. In total, four roles were established, those being the executive sponsor, as well as Chief, Business- and Technical Data stewards. Besides these roles the Data Quality Board was also established for corporation-wide DG definition and implementation monitoring. The additional influence factors are (1) performance strategy, (2) diversification breadth, (3) organization structure, (4) competitive strategy, (5) the degree of process harmonization, (6) degree of market regulation and (7) decision-making style (Weber et al, 2009, II).

Assessment: The Data Governance research contribution of Weber et al (2009, II) describes a structured research approach for collaborative development of a DG organization. Its only limitation is the lack of an explicit theoretical foundation.

Otto, B. (2011). Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers.

Otto (2011) focused on Data governance as a mean to ensure Data Quality and its management, however he also failed to state an explicit theoretical foundation for his research. He conducted a case study in this regard, including two major telecommunication companies. After describing the DQM and DG initiatives at the companies. The two companies were compared with regards to Data Governance along the lines of (1) organizational goals, (2) organizational form and (3) organizational transformation (Otto, 2011). Ultimately, while similarities in the design of Data governance were found with regards to goals, there is no one-size-fits-all approach. Sub-dimensions such as change measures and the transformation process vary even between companies from the same industry, therefore, Data Governance configuration is ultimately very dependent on external and internal factors (Otto, 2011). The issue of external validity of these findings is also addressed, with the conclusion that the results may be transferred to companies with complex IT as well as with a large customer base.

Assessment: Otto (2011) provides an interesting insight into the differences in Data Governance organization even in companies from the same industry as similar sizes. The conducted case studies and result statements appear to be sound, with the only exception being the common theme of the lack of an explicit theoretical foundation. Otto (2011) however aims to ultimately develop a well-developed theory on the organization of Data Governance, which might then act as the foundation of future research contributions regards DG.

Other Data Governance and DQM research contributions

Wende (2007) presents a Data governance framework, incorporating both Data Quality roles, decision areas and responsibilities, with a governance RACI matrix being proposed. Data Quality roles include e.g. the Chief (Data) steward, while decision areas include DQ strategy, DQ organization and DQ information system (Wende, 2007). Finally, the assignment of responsibilities refers to the assignment of the common RACI classification of (1) responsible, (2) accountable, (3) consulted and (4) informed (Wende, 2007).

This model for Data Governance and Data Quality Management is added upon by Weber et al (2008), who conducted a case study with regards to these issues. They identified a clear need for a formal Data governance model, as well as the steps in establishing this model. Besides that, there is also a need for the ability to carry out actions as result of this Data governance, as well as for the Data governance framework to be simple (Weber et al, 2008). Tying into this, Weber et al (2009, I) find the need for shared service center in their similar research regarding the organization of Data Quality Management. Weber (2009) similarly constructs a model of reference for Data governance, which includes among other things both internal and external conditions as well as a DQM characterization. This reference model can be used to plan the implementation of Data Quality Management by including five roles and two bodies that are responsible for Data Quality Management. Weber (2009) also include a design-object model that includes the six different views of strategy, leadership-system, organization, Data management processes, Data architecture and system support. It also includes both a function diagram for responsibility designation, as well as an approach model for the implementation of the Data governance reference model.

Liaw et al (2014) also investigated these two concepts with regards to their relevance in eHealth and healthcare. They found a need for an integration of Data Quality Management and Information governance, as well as proposing three different models for the organization of Data Quality Management and Data governance in healthcare providers. According to Liaw et al (2014), both Information governance and Data Quality Management have to be "good", in order to enable good decision-making. Their alignment is key, but it was found that the alignment between Data Quality Management and organizational objectives as part of an Information ecosystem, governed by an Information governance framework, is often lackluster (Liaw et al, 2014). This alignment between Data Quality Management within an Information governance framework is described as a nested system. The relationship between Data Quality Management and Data governance is also a topic in Kreis (2017), although in the context of Data migration rather than e-health. Both Data governance and Data Quality Management are identified as success factors for the successful migration of Data. Al-Ruithe et al (2019) added to this issue by analyzing the state of the art of Data governance, with Data Quality Management being one of the topics.

4.3.2 Research contributions related to DQM efficiency

Another concept identified in the research with regards to DQM is that of its efficiency with regards to the necessary investments, which was mentioned by three sources within the analyzed literature (see Table 13).

Research contribution	Goal	Foundation	Method
Heinrich & Klier (2006)	**Optimization of DQM regarding sales campaigns**	**No clear foundation is given**	**Mathematical modeling & Case study**
Baghi (2017)	Development of a Capability Model for DQM controlling	Resource-based-View	Design Science
Schäffer et al (2018)	Development of Model for Analysis and Calculation for the valuation of investments into DQM	No clear foundation is given	Design Science

Table 13 Research contributions related to DQM efficiency

Exemplarily, the research contribution of Heinrich & Klier (2006) is analyzed.

Heinrich, B., & Klier, M. (2006). Ein Optimierungsansatz für ein fortlaufendes Datenqualitätsmanagement und seine praktische Anwendung bei Kundenkampagnen

Heinrich & Klier (2006) tried to optimize Data Quality Management with regards to its practical use in regarding sales campaigns. Therefore, they adopted a mathematical model of the investment in DQM, utilizing mathematical modeling as their method. Their model was applied to a case study at a telecommunication company. Investment into DQM was found to be dependent on three factors. The first factor which has to be considered with regards to investments in DQM is the amount of Data, with the second being the percentage of base transactions (Heinrich & Klier, 2006). The third factor which was found to influence the efficient investment in DQM was the existing Data Quality level. Besides that, they also found that a high degree of Data Quality deterioration warrants higher investment in DQM, not as initially assumed lower investments. Another finding was that a low effectiveness of DQM measures did not lead to lower investments, but in fact required a higher degree of measure utilization (Heinrich & Klier, 2006).

Assessment: The Method for DQM investment optimization developed and tested by Heinrich & Klier (2006) provides an interesting and very formal approach towards the valuation of DQM. Its only apparent limitation is again the lack of an explicit theoretical foundation.

Other DQM efficiency research contributions

Data Quality Management efficiency and the need to justify investments into DQM were also addressed by Baghi (2017), who developed a capability model for DQM controlling.

One of the included case studies explicitly monitored investments into DQM. Schäfer & Beckmann (2018) developed a calculation model for Data Quality Management investments, differentiation between primary and secondary efficiency factors. Primary efficiency factors regarding DQM are Data Quality costs and technical value potential, while secondary efficiency factors refer to the attributes and influence factors of investments (Schäfer et al, 2018).

4.3.3 Research contributions related to DQM maturity

With regards to the systematic Management of Data Quality, its maturity is also an issue already hinted at with regards to Data governance transformation and prevalent in the reviewed literature. Five contributions were identified with regards to this topic (see Table 14).

Research contribution	**Goal**	**Foundation**	**Method**
Ryu et al (2006)	**Introduction of the DQM maturity model as a preferred maturity model for metadata management**	**No foundation is given**	**Survey**
Caballero et al (2008)	**Description of the foundations and structure of an Information Management Maturity Model**	**No foundation is given**	**No clear research method**
Hüner et al (2009)	Proposal of a reference model for CDQM maturity assessment.	No foundation is given	Design Science
Ofner et al (2009)	Development of a methodical approach the transformation of reference model into a company specific DQM maturity model.	No foundation is given	Design Science
Ofner et al (2013)	Presentation of a maturity model for Enterprise Data Quality Management	No foundation is given	Design Science

Table 14 Research contributions towards DQM maturity

Exemplarily, the research contributions of Ryu et al (2006) and Caballero et al (2008) are analyzed.

Ryu, K. S., Park, J. S., & Park, J. H. (2006). A Data Quality Management Maturity Model

Ryu et al (2006) introduce a DQM maturity model in the context of metadata management. The theoretical foundation of their research however remains unclear, with no theory being explicitly stated its foundation. As a research method, a two stage-survey is conducted for model verification. Their research model differentiated between four stages of Data Quality Management maturity, with the first one being the (1) initial phase. This is followed by the (2) defined and the (3) managed stages, concluding with the (4) optimized phase (Ryu et al, 2006). Besides a description of the individual stages, the usual issues of each stage are also described, as are the solutions for these. On the first stage of the survey interviews with CIO's from six different companies were used for the determination of Data management maturity levels. On the second stage, employees were divided into two groups and asked to compare DQ before and after Data Model management in group A. In Group B, they were asked to compare DQ before and after meta-data management. 46 employees were questioned in Group A and 73 in Group B. Overall, the survey verified the effectiveness of the proposed model as well as the research hypothesis. The first research hypotheses referred to the DQ and DQM maturity link, while the second referred to the greater difference between the third and the second stage than between the first and the second stage with regards to Data integration.

Assessment: Ryu et al (2006) provide an important research contributions with regards to DQM maturity. The validate their research via a two-stage survey, succeeding in proving effectiveness of the model as well as their hypotheses. The only limiting factor is again the lack of a sound theoretical foundation.

Hüner, K., Ofner, M., & Otto, B. (2009). Towards a maturity model for corporate Data Quality Management.

The development of a maturity model of DQM is the goal of Hüner et al (2009). For this they adopt a Design Science approach and conduct reference modeling. Their model of corporate Data Quality Management maturity consists of three dimensions. The first of these dimensions is the propagation of corporate Data Quality Management, which refers to the need for step-by-step implementation of DQM, as well as to the percentage of parts of the company which are already covered by DQM (Hüner et al, 2009). The second dimension of DQM maturity are corporate Data classes. These refer to the percentage of Data classes of the company, which are covered by DQM. Hüner et al's (2009) third and last dimension of

DQM maturity is that of the progress of practices and goal achievement. This dimension of DQM maturity refers to the need to monitor the progress of DQM implementation as well as to establish a general project management structure (Hüner et al, 2009). This maturity model is then applied to an electrical engineering company for evaluation. Ultimately, it was found that the model can be applied but requires further research regarding the requirements as well as regarding weights for particular goals (Hüner et al, 2009).

Assessment: The model developed by Hüner et al (2009) constitutes a well-structured research contribution towards Data Quality Management maturity. Only the lack of a clear research foundation limits its research quality.

Other DQM maturity research contributions

Caballero et al (2008) developed a framework for the management of Information Quality, which also includes an Information Quality management maturity model as one of its two components. This maturity model, IQM3, contains five maturity levels, each except for the first referring to a specific Information Quality management goal and having its own key process area. The first of these levels is the Initial level, which is followed by the Defined level (Caballero et al, 2008). The Defined level contains e.g. the critical process area of user requirement management. The third level, Integrated, contains the critical process area of risk and poor Information quality impact management, while the fourth level Quantitative Managed, contains e.g. IMP-measurement management. The fifth and final level is the Optimizing level and it contains among others the Causal-Analysis for Defect Prevention management area (Caballero et al, 2008).

Ofner et al (2009) add to this topic of DQM maturity by taking the already existing model of Hüner et al (2009) and accounting for company specific factors. They ultimately demonstrate the applicability of this model via two case studies. The model utilized by Ofner et al (2009) is the Corporate Data Quality Management Maturity (CDQ MM), containing six design areas as well as goals and practices. In order to assess the DQM maturity, a hierarchical approach is adopted, accounting for the different layers of the domain, design areas & goals and practices (Ofner et al, 2009). Ultimately, it is concluded in a number of assessment phases, subdivided by processes. These phases included the (1) requirement definition, (2) the adaption of a domain reference model as well as of (3) an assessment model. The last phase of Data Quality Management maturity assessment according to Ofner et al (2009) was the (4) preparation of the assessment. Their model also included five roles for the implementation of the maturity model, those of Sponsor, Data Steward, Business Data Steward, Process Owner and Process User (Ofner et al, 2009).

Ofner et al (2013) later revisit the issue with a more detailed and hierarchical Data Quality Management maturity model, this time including detailed practices and measures which are missing from previous models. It is based on the EFQM Excellence Model, but also integrates the context of the assessment (Ofner et al, 2013). It also emphasizes a stronger understanding of practices, assigning distinct methods and models to them. The last focus of this model is to allow for a company specific configuration (Ofner et al, 2013). It ultimately contains a wide variety of model constructs for the main domain as well as assessment, allowing for a great degree of customization and to account for company specific factors (Ofner et al, 2013).

4.3.4 Research contributions related to DQM frameworks

Another DQM concept identified in the research contributions were general DQM frameworks. Five contributions were identified with regards to this.

Research contribution	Goal	Foundation	Method
Lucas (2010, I)	**Development and Application of a Framework for DQM**	**No clear foundation is given**	**Case study**
Lucas (2010, II)	Understanding the nature and complexity of corporate DQ management	No clear foundation is given	Case study
Bargh et al (2015)	Proposal of a Framework for dynamic DQM	No clear foundation is given	Design Science
Frehe et al (2016)	Development of a Balanced Scorecard for DQM with regards to Big Data	No clear foundation is given	Design Science
Leadbetter et al (2020)	**Implementation of a DQM framework**	**No clear foundation is given**	**No clear method is described**

Table 15 Research contributions related to DQM frameworks

Exemplarily, the research contributions of Lucas (2010, I) and Leadbetter et al (2020) are analyzed.

Lucas (2010, I). Corporate Data Quality Management: From theory to practice
This research contributions aims at the development and application of a framework for Data Quality Management. However, it is based upon concepts such as Data Quality and Data Quality Management Maturity and not well-established theories. In order to achieve its goal, a case study approach is adopted. Lucas (2010, I) highlights the different areas and considerations towards Data Quality Management, such as Data Quality Assessment and DQM maturity. (Lucas, 2010, I). The observed company applied a limited bottom-up approach towards DQM, not using any established Data Quality Management methodologies and only a single Data Quality tool. Therefore, the observed company also lacked certain Data Quality Management roles and did not yet reach the proactive Data Quality Management maturity stage (Lucas, 2010, I).

Assessment: The research contribution of Lucas (2010, I) provides an interesting look into the case of a company on the reactive level of DQM and only a very limited bottom up DQM effort. However, the lack of a well-established theoretical foundation again decreases the overall quality, with it basing its research upon concepts such as DQ and DQM maturity.

Leadbetter, A., Carr, R., Flynn, S., Meaney, W., Moran, S., Bogan, Y., Brophy, L., Lyons, K., Stokes, D., & Thomas, R. (2020). Implementation of a Data Management Quality Management Framework at the Marine Institute, Ireland
The main goal of Leadbetter et al (2020) is the implantation of a DQM framework at an Irish Marine Institute. However, the underlaying foundation is limited to references regarding ISO standards. Besides this the research method remains unclear. Leadbetter et al (2020) describe several DQM roles during the implementation of DQM framework. These roles are (1) Data Owner, (2) Data Coordinator: Data Steward and (3) Data Protection Officer. The overall Framework contains the four key areas of Inputs, Planning, Operations and Output. These are supported by two additional areas of focus, Stakeholder Inputs and Performance evaluation. Besides these roles and areas of DQM at the Marine Institute an implementation pack was also developed, containing such elements as a Data Management plan guidelines and process flow examples (Leadbetter et al, 2020).

Assessment: Overall Leadbetter et al (2020) provides a very interesting insight into the practical the implementation of DQM at a small non-enterprise organization. However, its lack of both a distinct theoretical foundation and of a described method of framework development limits its overall quality. The overlap with Lucas (2010, II) is also concerning with regards to originality.

Other general DQM framework research contributions

Lucas (2010, II) ultimately only gives a more detailed description of Lucas (2010, I), providing more details with regards to the companies DQM initiative. Bargh et al (2015), on the other hand develop a framework for dynamic Data Quality Management. This framework consists of two partly overlapping areas, Data Quality Assessment, and Problem solving/Data Quality Improvement. Data Quality Assessment includes the identification of DQ attributes, the ranking of these attributes as well as their categorization. It also includes the problem registration, semantic field processing and the mapping of the problem to Data Quality attributes steps, which overlap with the problem-solving area. Problem solving then contains the steps of problem clustering, problem solving and problem severity measurement (Bargh et al, 2015). This Data Quality Management framework is then evaluated by Bargh et al (2015), confirming its usability for Data Quality Management.

Frehe et al (2016) also provide contribution towards the issue of Data Quality Management frameworks by developing a balanced scorecard in order to support systematic Data Quality Management in the context of big Data. Their balanced scorecard incorporated the four big Data related issues of Velocity, Veracity, Volume, and Variety, with the fifth issue of Value from Data being dependent of these four (Frehe et al, 2016). Each of these four main issues includes its own Data Quality dimensions, with such aspects as targets and metrics still needing to be defined by the applying company.

4.3.5 Research contributions related to the TDQM Framework

Another specific framework for DQM mentioned in the reviewed literature is the TDQM.

Research contribution	Goal	Foundation	Method
Wang (1998)	**Development of a methodology for Total Data Quality Management**	**No clear research foundation**	**No clear method is described**
Francisco et al (2017)	Comparison of TDQM and TIQM with regards to Customer Relationship Management	No clear research foundation	Comparative analysis

Table 16 Research contributions related to the TDQM Framework

Exemplarily for the two research contributions referring to this framework the one of Wang (1998) is analyzed.

Wang (1998). A product perspective on total Data Quality Management

The article by Wang (1998) aims at facilitating the implementation of an organization Data Quality policy by developing a methodology for Total Data Quality Management. Wang (1998) bases its research both in the TDQM cycle as well as in the concept of the Information product. However, the research contribution lacks a clear and explicit theoretical foundation, with both foundations lacking the nature of a full theory. Also, no clear method is stated in order to develop this TDQM methodology. Wang (1998) adopts an implicit Design Science approach, however lacking evaluation. Wang (1998) generally describes the process of applying the TDQM also as a circular pattern of IP definition, measurement, analysis, and improvement. Ultimately, this application of TDQM concludes in the improvement of the Information product, which according to Wang (1998) requires the identification of such key areas as the alignment between both the Information and workflow with their corresponding Information manufacturing system. In order to implement TDQM according to Wang (1998,) a company needs to follow four steps. Articulation of the Information Product in business terms is the first step from an organizational perspective of the TDQM according to Wang (1998). This is followed by the establishment of an IP team, as well as by teaching of IQ assessment and management skills. Ultimately, TDQM requires an organization to institutionalize continuous IP improvement.

Assessment: Wang (1998) forms an important cornerstone of the Total Data Quality Management framework, describing both the DQM process as well as how to implement it at an organization. However, from a research perspective, it suffers from its lack of both a clearly stated and well-established theoretical foundation as well as from the lack of a clear research method through which its results are derived. Although its general approach shares similarities with Design Science, this is neither formally expressed nor is an evaluation of the TDQM artifact conducted.

Other TDQM references

Another approach towards frameworks for Data Quality Management is provided by Francisco et al (2017). They discuss and compare the two frameworks of Total Data Quality Management and Total Information Quality Management with regards to their application to customer relationship management. On one hand Total Data Quality Management consists of four basic processes according to Francisco et al (2017), those being Definition, Measurement, Analysis, and Improvement. It focused on the root causes for Data Quality problems, seeing them as requirements for an improvement plan. Total Information Quality Management on the other hand assumes a Data Quality project as a process involving the entire organization (Francisco et al, 2017). It consists of the five steps of (1) meaning and structure assessment, (2) Information Quality assessment, (3) measurement of costs for bad Data Quality, (4) Data reengineering and cleaning and (5) Information generation improvement. It also includes a sixth process, containing tasks and activities relating to the five previously mentioned processes and referred to as an action plan by Francisco et al (2017). These two frameworks were compared with regards to three criteria, those being difficulty of implementation, efficiency, and completeness. Ultimately, it was found that TDQM allows for a more incremental approach towards Data Quality Management, limiting overall investment. However, TDQM excludes meta-data as well as failing to address the issue of costs through poor Data Quality, possibly allowing for poor efficiency. TIQM on the other hand is very suited towards meta-data and Data Reengineering, but more general in other regards and requires the participation of at least one expert in its methodology (Francisco et al, 2017).

4.3.6 Research contributions related to the CDQM Framework

Three references referred to the CDQM as described by Österle & Otto (2016).

Research contribution	Goal	Foundation	Method
Otto et al (2007)	**Development of a framework for Corporate Data Quality Management**	**No clear foundation is given**	**Design Science**
Otto & Hinderer (2009)	Provision of an architecture design for DQM in supplier controlling	No clear foundation is given	Design Science
Österle & Otto (2016)	Provision of an overview over Corporate Data Quality Management	No clear foundation is given	No clear method is described

Table 17 Research contributions related to the CDQM Framework

Exemplarily, the research contribution of Otto et al (2007) is analyzed.

Otto, B., Wende, K., Schmidt, A., & Osl, P (2007). Towards a framework for corporate Data Quality Management

Otto et al (2007) aim at developing a framework for Data Quality Management, the Corporate Data Quality Management (CDQM) framework. No clear and explicit theoretical foundation is given. For the development of the framework an implicit Design Science approach is adopted. Their framework is based upon the method of business reengineering and its three areas of Corporate Data Quality (CDQ) strategy, CDQ organization, and CDQ architecture. These three areas can be subdivided into the layers of Governance and Execution, leading to a total of six CDQ practices. These are (1) Development of a CDQ strategy, (2) Design of the CDQ organization, (3) Design of the CDQ architecture, (4) Communication and Control of the CDQ strategy, (5) Execution and Monitoring of CDQ processes, and (6) Operation and Maintenance of the CDQ architecture. The DQM framework established by Otto et al (2007) was evaluated in a first evaluation cycle, but still lacks formal verification.

Assessment: Otto et al (2007) provide a framework to corporate Data Quality Management. The usefulness of that framework is however limited by the lack of formal verification as well as its unclear theoretical foundation.

Other CDQM references

Another contribution towards Data Quality Management frameworks is provided by Otto & Hinderer (2009). They applied the already mentioned CDQM framework to the context of vendor controlling. For this, they applied it to four different case examples, deriving an architecture draft from these. They also identified several influence factors on the applicability of CDQ, such as the degree of centralization.

Another contribution towards the CDQM is also provided by Österle & Otto (2016). They formulate it as framework for master Data Quality Management and discuss it on the premise of ten case studies, some of which have been discussed before. The master Data Quality Management framework of Österle & Otto (2016) consists of three layers, those of DQM strategy, DQM organization and DQM Information system, with a total of six design areas as well as individual types of results.

(1) **Data Quality Management strategy:** This layer refers to the alignment of the Data Quality Management with the business goals (Österle & Otto, 2016).
(2) **Data Quality Management organization:** The organizational layer of Data Quality Management according to Österle & Otto (2016) is best described by its three design

areas: (1) leadership system for Data Quality Management, (2) Data Quality Management organization and (3) Data Quality Management methods and processes as a design area (Österle & Otto, 2016).

(3) Data Quality Management information system: This last layer of Data Quality Management can be described by its two areas of Data Quality Management architecture and Data Quality Management applications (Österle & Otto, 2016).

4.3.7 Research contributions related to Process-Driven-Data Quality Management

Three research contributions referred to the topic of Process-Driven-Data Quality Management (PDDQM) (see Table 18).

Research contribution	Goal	Foundation	Method
Grimmer & Hinrichs (2001)	**Illustrating the means by which DQ problems are addressed in practice**	**No clear foundation is given**	**Case study**
Glowalla & Sunyaev (2013)	Highlighting the two options of within-model integration and across-model integration to integrate Data Quality into existing process models	No clear foundation is given	Literature review
Glowalla & Sunyaev (2014)	Provision of a synthesis of the possible applications of process modeling languages for PDDQM	No clear foundation is given	Literature review

Table 18 Research contributions related to Process-Driven-Data Quality Management

Exemplarily, the research contribution Grimmer & Hinrichs (2001) is analyzed.

Grimmer & Hinrichs (2001). A Methodological Approach to Data Quality Management Supported by Data Mining

Grimmer & Hinrichs (2001) aim a providing a process-oriented approach towards DQM. They do not establish a well formulated and explicitly stated theoretical foundation for their research, limiting themselves to basing it on concepts such as DQM and standards such as ISO9001. In order to achieve the research goals, a process model for Data integration is proposed and applied to a case in the automotive industry. The process model consists of the ten steps of (1) Unification of Representation, (2) Statistical Process control, (3) Domain-specific consistency checking, (4) inquiries & postponing, (5) record linkage, (6) merging, (7) Quality Measurement and Analysis, (8) control of nonconforming Data Products and Quality Improvement, (9) Data Product release and target area update and (10) Analysis of customer feedback & retraction of Data products (Grimmer & Hinrichs, 2001). It forms the basis of their Data Quality Management System, with selected steps being implemented in the case study.

Assessment: The research contribution of Grimmer & Hinrichs (2001) aims at providing a process-oriented approach towards DQM via its process model which is implemented in the case study at an automotive company. In this goal it succeeds. Besides the lack of a well-developed theory as a research foundation there are no apparent issues with regards to this research contribution.

Other PDQM references

Glowalla & Sunyaev (2013) further add to the topic of Process-Driven Data Quality Management by developing a research framework with regards to the application of process modeling languages and the selection of integration approaches. This issue is revisited by Glowalla & Sunyaev (2014), who derive requirements for the application of process modelling languages for PDDQM.

4.3.8 Research contributions related to the Information MAP

Two of the identified theoretical research contributions also reference the Information MAP (IMAP).

Research contribution	**Goal**	**Foundation**	**Method**
Shankaranarayan et al (2003)	Presentation of a framework for DQM in dynamic decision-making environments	Information Product	Unclear method
Shankaranarayan & Cai (2006)	**Proposal of a decision-support framework**	**Information Product**	**Unclear method**

Table 19 Research contributions related to the Information MAP

Exemplarily, the contribution of Shankaranarayanan & Cai (2006) is analyzed.

Shankaranarayanan & Cai (2006). Supporting Data Quality Management in decision-making

Shankaranarayanan & Cai (2006) aim at providing a framework for decision support. They base their research in the concept of the Information product, however this lacks the nature of a clearly developed causal theory. In order to develop this framework a Design Science approach is used. They also utilize the Information Product Map (IPMAP) as a means to represent the manufacture of an Information product. An IPMAP consists of such elements as Data source blocks, Data processing blocks, Data storage blocks, inspection blocks and Data sink blocks. It also includes such elements as the Information system boundaries and the business/organizational boundaries (Shankaranarayanan & Cai, 2006).

Assessment: Overall Shankaranarayanan & Cai (2006) provides an interesting yet flawed insight into this issue, with it being limited by both the lack of an evaluation of the developed framework as well as of a clear theoretical foundation.

Other IMAP references

A Data Quality Management tool/framework is provided by Shankaranarayanan et al (2003). Both IMAP constructs and capabilities for Data Quality Management related to the Information product are described. Capabilities include e.g. time-to-delivery, with a virtual business environment being proposed to support dynamic decision making. This IPMAP is then utilized for a framework to evaluate the Information Quality dimension of completeness with regards to an Information product, the requirements planning report (Shankaranarayanan & Cai, 2006). The developed IPMAP is then used for visualization for the developed decision support tool, IPView. This decision support tool can then be used with regards to establish Total Data Quality Management as well as to evaluate Data Quality dimensions (Shankaranarayanan & Cai, 2006).

4.3.9 Research contributions related to Master Data Quality Management

Another concept identified in the systematic review is that of Master Data Quality Management (MDQM), with two references referring to this topic.

Research contribution	Goal	Foundation	Method
Otto & Hüner (2009)	Development of a reference architecture for Master Data Management	No clear foundation is given	Design Science & Case study
Otto et al (2012)	**Development of a reference architecture for Master Data Management**	**No clear foundation is given**	**Design Science & Case study**

Table 20 Research contributions related to Master Data Quality Management

Exemplary on Otto et al (2012) these are analyzed.

Otto et al (2012). Toward a functional reference model for master Data Quality Management

Otto et al (2012) aim at answering the question which functionality a system for MDQM has to provide by developing a functional reference architecture. They references the concepts of Master Data Management as well as Master Data Quality Management but fail to establish an explicit theoretical foundation for their research. In order to develop this reference architecture, a Design Science approach was adopted, with the research process being conducted in six steps. Ultimately, this reference model includes six function groups with a total of 19 individual functions. The six main function groups are (1) MD Lifecycle Management, (2) Meta Data Management & Master Data Modeling, (3) DQ Assurance, (4) MD Integration, (5) Cross functions and (6) administration. This reference model provided by Otto et al (2012) was then applied to a case study, which resulted in the four steps of (1) preparation, (2) assessment, (3) as-is- and (4) as-to-be analysis.

Assessment: Overall Otto et al (2012) provide an interesting contribution towards Master Data Quality Management, as well as useful artifact in their reference architecture. The reference architecture was evaluated in terms of practical applicability via the case study. Therefore, the only limiting factor is the lack of an explicit theoretical foundation.

Other MDQM references

Another contribution towards the concept of Master Data Quality Management is provided by Otto & Hüner (2009), by proofing a functional reference architecture for master Data management, master Data lifecycle management and master Data Quality Management. Their master Data Quality Management framework includes the three functional areas of Data analysis, Data enrichment and Data cleansing.

4.3.10 Research contributions towards DQM strategy

Besides Data Quality Management frameworks, several identified references also referred to specific aspects of DQM, which can be organized along the lines of the Corporate Data Quality Management framework provided by Österle & Otto (2016). One of these is the DQM strategy.

Research contribution	Goal	Foundation	Method
Falge et al (2013)	Presentation of a method for development and implementation of a CDQM strategy	No clear theoretical foundation	Design Science
Falge (2014)	Presentation of a method for development and implementation of a CDQM strategy	No clear theoretical foundation	Design Science

Table 21 Research contributions towards DQM strategy

Exemplarily, Falge (2013) is analyzed.

Falge (2013). Towards a Strategy Design Method for Corporate Data Quality Management

Falge (2013) wants to address the need for a DQM strategy by developing a method its development. The paper builds heavily upon existing concepts of DQM as well as Data Governance but fails to explicitly establish one clear and explicit theoretical foundation. It is developed in a collaborative Design Science Approach. The method contains the four strategy development phases of (1) Analysis, (2) Strategy Development, (3) Justification, (4) Implementation and Monitoring. The method was also evaluated in focus group interviews and participatory case studies.

Assessment: The strategy development method developed by Falge (2013) provides an evaluated artifact for this purpose, without clear indications for the violation of research criteria. The only noteworthy limitation is the lack of an explicitly formulated theoretical foundation.

Other DQM strategy references

The overall efficiency of DQM was also a research topic in Falge (2014), which describes the design process of method for the development of strategies regarding Data Quality Management in greater detail. Part of this method was an efficiency analysis, which was used both to justify investment in DQM as well as to monitor costs. It was preceded by the analysis and strategy development phases and formed the penultimate phase of the developed method before the implementation phase. This efficiency analysis included a total of three steps, with the identification of the effect sequence being followed by the establishment of transparency with regards to costs and value (Falge, 2014). The last step of this efficiency analysis was the investment calculation. Falge (2014) applies the already discussed method for strategy development to five case studies at a variety of companies. She also develops a meta-model for DQM, concluding by evaluating both the method for DQM strategy development and the meta-model for DQM (Falge, 2014).

4.3.11 Research contributions related to DQM system

Besides the two layers of DQM strategy and -organization, the third layer of DQM, the system as described by Österle & Otto (2016) was also among the research topics in the identified references.

Research contribution	**Goal**	**Foundation**	**Method**
Madnick et al (2003)	**Outlining a technical approach to a corporate householding knowledge processor to solve entity aggregation**	**No clear foundation is given**	**No clear method is described**
Winter et al (2003)	Analysis and description of DQM from the technical and organizational perspective at Credit Suisse	No clear foundation is given	No clear method is described
Schmidt et al (2010)	Description of a method for modeling of Data	No clear foundation is given	Case study
Westin & Sein (2015)	**Development and implementation of a DQ/IQ assessment tool**	**No clear foundation is given**	**Action research**
Schäffer & Stelzer (2017)	Identification and Assessment of coordination mechanisms and tools used to facilitate product Information sharing	No clear foundation is given	Case study

Table 22 Research contributions related to DQM system

Exemplarily, Madnick et al (2003) and Westin & Sein (2015) are analyzed.

Madnick, S., Wang, R., & Xian, X. (2003). The design and implementation of a corporate householding knowledge processor to improve Data Quality

Madnick et al (2003) aim at addressing DQ problems in corporate householding. They base their research in corporate householding as well as DQ, both of which however do not meet the requirement of being a full theory. In order to improve DQ in corporate householding a knowledge processor (CHKP) is developed via a Design Science approach. CHKP can be used to counter a specific Data Quality problem, Data aggregation, and is based on the Context Interchange Technology (COIN). This consists of three component processes, client, server, and mediator processes, with Madnick et al (2003) concluding in developing and implementing a corporate householding query processor based on it. Therefore, the designed processor is evaluated.

Assessment: Madnick et al (2003) develop and successfully evaluate a usable IT artifact in their research contribution. Its only limiting factor is the lack of an explicitly described theoretical foundation and research method for the development of the knowledge processor.

Westin, S., & Sein, M. (2015). The Design and Emergence of a Data/Information Quality System

Westin & Sein (2015) aim at developing a Data/Information Quality assessment tool IQS. However, they fail to describe a clear theory as the foundation of their research. The tool was developed over to action design research cycles in the construction industry, with different objectives in each phase. Overall the establish a revised set of five design principles for tool development, which are (1) allowing for inconsistencies, (2) allowing for incompleteness, (3) allowing for lack of logical coherence (4) phase-based reporting, and (5) parking of errors.

Assessment: Overall Westin & Sein describe an interesting example of the development of a tool for DQM. The only limiting factor is the lack of a clear theoretical foundation.

Other DQM system references

Winter et al (2003) describe a Data Quality module as a part of the meta-data administration. It includes elements such as Data Quality rules, Data Quality statements based on these rules and a logfile for Data Quality problem tracking in case a Data Quality rule is broken (Winter et al, 2003). The model discussed by them consists of the four phases of Planning, Doing, Checking & Acting, and is later applied to a case study at a major bank. Ultimately, Winter et al (2003) conclude by describing both the technical and organizational aspects of Data

Quality Management, with the organizational side referring to roles and responsibilities for Data Quality problems, which can be either triggered by rules or Data Quality incidents.

Schmidt et al (2010) provide another limited contribution towards this issue by describing how the observed company handled Data Quality Management related issues with its IT architecture. In the observed company the IT architecture was not the responsibility of the DQM function but rather the corresponding corporate and business functions, an issue which was tackled by devising a Data map under the involvement of DQM functions. This map could then be used to analyze the use of Data within the company with regards to which applications accessed which Data objects (Schmidt et al, 2010). Besides this/that, four layers of corporate architecture are described, them being business process architecture, professional architecture, system architecture and technical architecture. This included the identification of responsible roles for specific Data categories such as customer Data, with several roles such as Data architects and managers for each Data category (Schmidt et al, 2010).

Schäffer & Stelzer (2017) also add to this area by assessing tools for product master Data Quality coordination in three case studies. In total eight tools for product Data Quality coordination are analyzed, both in terms of strengths and weaknesses. After describing these tools for the coordination product Information Quality Schäffer & Stelzer (2017) conclude by describing the implications of the case study results.

4.4 Analysis and assessment of DQ/DQM research-oriented contributions

Besides the already discussed topics with regards to Data Quality and Data Quality Management, eight of the reviewed contributions also focused on Data Quality and Data Quality Management research, similar to this paper.

Research contribution	Goal	Foundation	Method
Wang et al (1995)	**Categorizing DQ research contributions via a framework**	**No clear foundation besides ISO9000 is given**	**Literature review**
Madnick et al (2009)	**Categorize DQ research contributions via a framework**	**No clear foundation is given**	**No clear method is given**

Jaya et al (2017)	Highlighting issues in Data Quality research and discussion of research opportunities	No clear foundation is given	Literature review
Nurminen (2017)	Definition and assessment of DQ as well as of current issues in Data Quality Management.	No clear foundation is given	Literature review
Shankaranarayanan & Blake (2017)	Identification of core topics and themes which define of Data Quality research	No clear foundation is given	Latent Semantic Analysis
Houston et al (2018)	Development of standard Data Quality monitoring procedures to ensure Data integrity	No clear foundation is given	Survey
Sautter et al (2018)	Proposal of a definition of the term Data excellence	No clear foundation is given	Literature review & Case study
Jaya et al (2019)	Explanation of the landscape of Data Quality and identification of research gaps	No clear foundation is given	Literature review

Table 23 Analysis and assessment of DQ/DQM research-oriented contributions

Exemplary by Wang et al (1995) and Madnick et al (2009) these are analyzed.

Wang, R. Y., Storey, V. C., & Firth, C. P. A framework for analysis of Data Quality research

Wang et al (1995) aim to categorize DQ research contributions by developing a framework for them. However, while referring to the ISO9000 standard as the basis of their research, they do not base their research upon a well-developed theory. They employ literature review as a method as well as an implicit design Science approach to develop a framework for DQ research. This framework consists of the seven elements of (1) management responsibilities, (2) operation and assurance costs, (3) research and development, (4) production, (5) distribution, (6) personnel management, and (7) legal function. This framework is then applied to a number of Data Quality articles and evaluated by categorizing these research contributions in the DQ context.

Assessment: Overall, the research contribution of Wang et al (1995) can be seen as providing a good contribution for DQ and DQM research. The framework is evaluated by application to a number of research articles, succeeding a good categorization of them. The limiting factors are that the theoretical foundation of Wang et al (1995) remains somewhat unclear, only loosely referring to the ISO9000 as a basis for their framework, as well as the lack of an explicit method.

Madnick, S. E., Wang, R. Y., Lee, Y. W., & Zhu, H. (2009). Overview and framework for Data and Information Quality research.

Madnick et al (2009) also aim at providing an overview as well as a framework for Data Quality research. However, both the theorical foundation of their research contributions as well as their method remains unmentioned. The framework consists of the four Data Quality research topics of (1) Data Quality impact, (2) Database related technical solutions for Data Quality, (3) Data Quality in the context of computer science and IT as well as (4) Data Quality in curation (Madnick et al, 2009). It also highlights fourteen prevalent methods in Data Quality research, among them action research and case study, but also econometrics and mathematical modelling.

Assessment: Madnick et al (2009) provide a useful contributions towards the organization of DQ research contributions. Their framework may be applied to categorize them along the lines of their topics as well as utilized methods. However, as both its own theoretical foundation as well as its research methods are not described, it suffers in terms of research quality.

Other DQ/DQM research references

Jaya et al (2017) also conducted a literature review who identified several methods for Data Quality assessment such as cell level tagging and control matrices. Nurminen (2017) similarly conducted a literature review in order to identify issues and topics in DQM research. Shankaranarayanan & Blake (2017) also conducted a software-based literature review, identifying eight overall topics in Data Quality research, each with corresponding themes. These topics include e.g. contextual assessment of DQ as well as DQM for networked Data and are also linked to corresponding Data Quality dimensions (Shankaranarayanan & Blake, 2017).

Another contribution towards this is provided by Houston et al (2018), who survey the prevalence of DQM in clinical research sites and found e.g. that only half of the surveyed research sites had a Data management plan, but that some kind of DQ related training was prevalent at every site. Sautter et al (2018) argue for a broader concept of Data Quality and its management than the ones previously discussed, referring to it as Data excellence, which is based on the Data suitability. This refers to organizational Data challenges, for which the four challenge dimensions of (1) operational excellence, (2) legal challenges, (3) Data Management process quality/maturity and (4) Data Quality (Sautter et al, 2018). Jaya et al (2019), again distinguish between different Data Quality research topics, in this case between (1) Data Quality impact, (2) Data Quality in computer science and (3) Database technical solutions. Also, different types of research as well as the different research methods are used in Data Quality research

4.5 Research contributions summary

In total the analyzed research contributions cover many different topics in the context of Data Quality and Data Quality Management and aim towards a wide variety of different research goals. In order to achieve this the utilize a wide variety of research methods, both implicitly and explicitly. Interestingly the research methods of design science and case study are especially prevalent. Design science is often used in the reviewed literature in order to develop some artifact, such as a framework, model, or method. Case study on the other hand is often used in order to assess how DQM is organized and conducted in practice, as well as to evaluate already existing or freshly developed artifacts. Both of these as well as the method of action research are also often used in the context of research contribution of the Competence Center Data Quality Management (CC DQM). This long-term project provides many of the reviewed references, besides the TDQM project of Richard Wang and others.

Another apparent fact is the lack of explicit and well-developed theoretical foundations. While many references either base their research in certain concepts or standards, such as Total Quality Management or ISO9000, the resource-based-view represents the only “real” theory. This however fits with the heavy use of Design Science research, which as stated in Chapter 4.1.2.3 can be utilized in context with a lack of clearly developed theoretical foundations.

5 Data Quality Management in Practice

Another issue of importance with regards to Data Quality Management is that of its practical design in operating organizations. Differences between types of organizations and industries of operation are also of interest. Over the course of the systematic review, 27 different companies were identified.

5.1 Data Quality Management in different industries

The expression of DQM in different industries, based upon the identified research contributions and practice sources is of also great interest. Both the case studies and action research projects were conducted in a wide variety of different industries, although with considerable overlap in companies in the context of the CC DQM. Overall, the distinction can be made between seven industries mentioned in the reviewed literature and practice reports. Besides these seven common industries, several companies from industries which are only mentioned a single time are also analyzed. These companies are analyzed in terms of (1) DQM practice at that company and (2) foundation of the companies DQM practice in the research contributions.

5.1.1 Telecommunications

Organization	**References**	**Source types**
British Telecom	Weber (2009), Otto (2011)	Research contribution
Deutsche Telekom	Schmidt et al (2010), Otto (2011)	Research contribution
MyTelecom (Anonymous)	Lucas (2010, I), Lucas (2010, II)	Research contribution
TelCo (Anonymous)	Falge (2014)	Research contribution

Table 24 Telecommunications companies

5.1.1.1 British Telecom

At British Telecom (BT) Data Quality is Management focused on Master Data such as customer Information and entry point. A DQM initiative became necessary due to heterogenous IT systems and the lack of a corporate-wide understanding of processes. This DQM initiative took the form of the Information Management Program, which focused on improving Data Quality in legacy system, assuring DQ in new system and accelerating the speed of migration from old to new systems (Weber, 2009). The central element of the project was the Information policy, a DQ strategy, while the IM forum was tasked with coordination between business and IT. This forum included both senior executives as well as technical specialists. Technical knowledge was provided by the IM-Team, a center of excellence, while the specialized projects derived from the Information policy were conducted by a designated separate team. During this program, each department was designated an Information Manager. Information Manager at BT are charged with assuring DQ in their department and communicating business requirements to the IM forum. Overall, the project was deemed successful, with a value of 700 Mio GBP being derived from it over its lifetime. The DQM organization was also continued in three separate function, dealing with Master Data Management, further DQM projects and offering DQM consulting services for customers.

Overall BT’s approach towards DQM can be seen as highly successful, with a quantifiable net gain as well as the establishment of a new capability that can even be offered to others for increased value from DQM. It is also well funded theoretically, with it being based mostly in the establishment of Data Governance roles and responsibilities as described by Wende (2007), with the IM forum as a supervisory board and the Information managers filling as similar role as business Data Stewards. It also incorporates the CDQ strategy development practice from the CDQM in the Information policy (Otto et al, 2007).

5.1.1.2 Deutsche Telekom

In the case of Deutsche Telekom (DT), a dedicated DQM initiative became necessary due to the merger of two departments, for which the need became apparent to assure the consistency of customer Master Data in order to offer combined products (Schmidt et al, 2010). Therefore, it was decided to implement a dedicated DQM function, with business requirements towards Data being handled by the Marketing and Quality Management department and the technical implementation, as well as the Design of concepts for DQM being handled by the department of Master Data Management. Master Data Management was therefore responsible for the development of corporate policy for DQM, DQ Assessment and for establishing corporate standards for Data. Besides this the, DQM initiative of Deutsche Telekom (DT) also included the assignment of roles and responsibilities for business Data objects as well as the definition of an IT architecture model as a basis for the modeling of Data. The value derived from this Data modeling was stated as up to 1% of the overall IT Budget.

Overall, DQM at the Deutsche Telekom provides a detailed process for implementing DQM and assuring DQ via the means of both Data Governance and the implementation of dedicated DQM functions. It is also theoretically founded via its utilization of Data Governance roles and responsibilities as described by Weber et al (2008) as well as via elements of the CDQM framework of Österle & Otto (2016).

5.1.1.3 MyTelecom (A)

Lucas (2010, I & II) presented the case another telecommunication company with the fictitious name of MyTelecom (A). At MyTelecom, DQM implementation took the form of a Data Governance Initiative, which focused on the operational level on the improvement of Master Data Quality such as customer names and addresses. Responsible for the DQM initiative is the IT department, however MyTelecom lacks any form of corporate Data policy or MDM, therefore its approach was very bottom-up and focused upon one particular Dataset. Overall, the DQM at MyTelecom is described as very reactive, not utilizing Data stewards or any established DQM methodology such as TDQM.

Overall, the case of MyTelecom provides an important example for a company whose approach towards DQM lack theoretical foundation. Their limitation to one Dataset, the lack of established DQM roles and responsibilities hints at a low level of DQM maturity. This is emphasized by the lack of organizational functions as well as lack of a DQM methodology and ultimately limits their success in assuring Data Quality.

5.1.1.4 TelCo

Another telecommunication company with a fictitious name is presented by Falge (2014), TelCo. At TelCo, DQM focused on the provision of high-quality Master Data for the offering of customer individual products. For this, its DQM included both Data Governance as well as corporation-wide understanding of process and partial automation via a workflow management system. TelCo also conducted a project for development of a DQM strategy via the method of Falge (2014), utilization three of its phases in (1) Analysis, (2) Strategy development and (4) implementation of the designed strategy.

Overall, DQM at TelCo seems to be both well organized as well as theoretically founded. It utilizes both Data Governance as well as having created a corporation wide understanding of processes and a partially automated system for them. The application of the CDQ strategy development method of Falge (2014) also establishes a clear foundation of its DQM on the strategy layer of Corporate Data Quality Management.

5.1.2 Healthcare and Pharmaceuticals

Organization	**References**	**Source types**
B. Braun	Weber (2009)	Research contribution
Johnson & Johnson	Österle & Otto (2016)	Research contribution

Table 25 Healthcare and Pharmaceutical companies

5.1.2.1 B. Braun

The healthcare company of B. Braun as described by Weber (2009) required a DQM initiative due to its diverse IT landscape and decentralized organizational structure. It focused on improving its Master Data Quality via a project called Central Data Master Server (CMS), which was conducted in the context of SAP consolidation project. CMS had the goals of (1) implementing a centralized Master Data system, (2) definition of global standards for Data, (3) definition and implementation of centralized processes for Master Data Maintenance and (4) partially automated Workflows. The organizational function for Master Data Management at B. Braun was the Central Material Master Agency, a part of the supply chain function. A specific focus was put on the points of transfer for Master Data, which are responsible for one kind of material Master Data. A difference was made between local points of transfer, whose Master Data refers to products which are only relevant for one subsidiary and not sold, and global points of transfer, whose Master Data refers to sold and often used material. An overall Governance structure was designed in this regard, as well as yearly meetings between the global transfer points scheduled. B. Braun also did not conduct prior value analysis, however the project succeeded in meeting its goals.

Overall, DQM at B. Braun as describe by Weber (2009) is very specific towards Material Master Data. In this, it provides an important example for the practical organization of DQM via Data Governance, with the transfer points assuming roles similar to those of Data stewards. Therefore, B. Braun's approach towards DQM is founded in such theoretical contributions as Wende (2007)

5.1.2.2 Johnson & Johnson

The fortune 500 company of Johnson & Johnson as described by Österle & Otto (2016) suffered from decentralized and heterogenous processes as well as from the lack of a corporation-wide understanding of the business objects. This led to various DQ issues, which Johnson & Johnson aimed at addressing via an initiative for Data Governance. For this, a corporate function for Master Data Management was implemented, one with central authority with regards to Data Quality. This function assigned roles and responsibilities for specific Data classes, initially based upon departments but later cross-sectoral. These are added to by both a steering committee and a yearly summit for Master Data. Besides this department, Johnson & Johnson also implemented organization-wide processes for Data Management. The DQM function also developed specific systems for assurance of Data Quality.

Overall, the case of Johnson & Johnson provides another example of the implementation of DQM functionalities via Data Governance. It implemented roles and responsibilities as well as a steering committee for DQM and is therefore very well founded in such research contributions as Weber et al (2008). Johnsons & Johnson's DQM also incorporates elements of the CDQM on the organizational layer via its design of new process of Data Management (Österle & Otto, 2016).

5.1.3 Agriculture

Organization	References	Source types
BayerCropScience	Weber (2009), Falge (2014), Österle & Otto (2016)	Research contribution
Syngenta	Österle & Otto (2016)	Research contribution

Table 26 Agriculture

5.1.3.1 BayerCropScience

BayerCropScience aimed at harmonizing its IT and application landscape. During this initiative, a number of Data Quality related problems became apparent, for which the root causes were identified. BayerCropScience therefore aimed at increasing Data Quality. For this, a Data Quality Cockpit software was devised and implemented in order to measure and visualize DQ. BayerCropScience also aimed at providing a consciousness for Data Quality as well as increasing transparency and identifying necessary actions. Besides this DQM Cockpit, Data Quality targets were also implemented as goals for employees, with the regional heads being responsible for Data Quality and having a 97% DQ level as a yearly goal. Regional heads were also responsible for assigning Data coordinators, employees which are tasked with establishing DQ roles and responsibilities within that country and with directing measures for DQ improvement. As a basis for the calculation of DQ levels, business rules were utilized, which are combined into a Data Quality Indicator. Another special system was also implemented in order to validate business rules.

Overall BayerCropScience's approach towards DQM is very extensive and heavily based in both Data Governance via role and responsibility assignment and the CDQM in its use of business rules and usage of dedicated DQM system (Wende, 2007; Österle & Otto, 2016).

5.1.3.2 Syngenta

The Data and Information Management initiative of Syngenta was one of three strategic initiatives of this company. This initiative aimed at consolidating the previously decentralized master Data resources under a singular leadership, as well as proving a standardized process for the provision of high-quality Master Data. It also provided a new centralized tool for Master Data maintenance. For this, a new Data stewardship organization was devised, including six roles for master Data Management such as the Master Data Management or the regional Data Steward. Many of the master Data tasks were outsourced, e.g. the Data cleaning. Due to this initiative, Syngenta was able to improve Data cleaning processes as well as standardize processes, although it also had to be assured that proprietary Information remained confidential and secured with regards to the external service providers.

Overall, this DQM initiative at Syngenta is again based both in Data Governance akin to Weber et al (2008) and the Corporate Data Quality Management framework of Österle & Otto (2016). It assigned both a new stewardship organization as well as a dedicated system for Data maintenance of the system layer of the CDQM.

5.1.4 Consumer goods

Organization	References	Source types
Beiersdorf	(Hüner et al, 2011), Österle & Otto (2016)	Research contribution
BayerConsumerCare	Falge (2014)	Research contribution

Table 27 Consumer goods companies

5.1.4.1 Beiersdorf

Beiersdorf tried to improve their intercompany product data sharing and therefore also focusing on Master Data (Hüner et al, 2011). Their previously diversified and spread out master Data tasks and responsibilities where unified in a central Product Lifecycle Management system, operated by the shared series department. This system included a master Data workbench for cleaning. Added to this are regional Business Data Stewards for each country, as well as one person from the marketing department being responsible for each product line. Besides this, a DQ defect and identification and monitoring project was conducted, leading to a number of DQ metrics which are evaluated monthly. A follow up project was also planned in order to devise and implement a system for these DQ metrics (Hüner et al, 2011).

Overall, Beiersdorf describes the case of an extensive implementation of Data Quality Management based on the Data Governance. Its establishment of roles for DQM such as the Regional and Product-Line Data Stewards ties into research contributions such as Weber et al (2008)

5.1.4.2 BayerConsumerCare

BayerConsumerCare, as described by Falge (2014) used a centralized SAP system and possessed a central Master Data Management for its sellable products as well as decentralized MDM for non-sellable products. However, the value of the DQM measures was not assessed, therefore BayerConsumerCare conducted a project in this regard, with the goals of identifying cost and quality potentials, designing an action plan, and conducting a value assessment. This resulted in the identification of cost intensive master Data management processes, based on this, short-as well as mid- and long-term solutions were devised. Short term solutions included such actions as making important attributes which generate a material number mandatory field, as well as using standard values for false Data entries. Both these short-term actions led to a decrease in process time and an overall DQ improvement of 33% (Falge, 2014). The mid-term solution on the other hand aimed at allowing the Master Data Managers quick insights into necessary changes, while the long-term solution focused on limiting Master Data involvement by establishing automated plausibility checks.

Overall, BayerConsumerCare's initial approach is only loosely based on reviewed theories. Its application of the strategy design method by Falge (2014) however is well founded in terms of theory.

5.1.5 Finance and Insurance

Organization	References	Source types
Allianz	Österle & Otto (2016)	Research contribution
Sparkassen Finanzgruppe	Finanz Informatik (2020, I), Finanz Informatik (2020, II)	Practice report
Credit Suisse	Helfert & Hermann (2002), Winter et al (2003)	Research contribution
Barmenia	ACT IT-Consulting & Services GmbH. (2011)	Practice report

Table 28 Finance and Insurance companies

5.1.5.1 Allianz

At Allianz, external pressure by government entities led to a need for the definition of dedicated DQM roles and responsibilities as well as their implementation. This was done in Allianz's Solvency II project, it founding a Data Governance team and having the three goals of the definition of DG requirements, definition of processes and actions in this regard and of designing action for DQ measurement and monitoring. Based upon these goals, an action plan was devised, with such actions as e.g. the establishment of a corporate policy for DQM and the implementation of a DQM system. A DMAIC cycle was used for process design, with the processes being based on quality management. As a result of the Solvency II project, DQ requirements were made by the users of the Data, while three DQ metrics referring to completeness, accuracy and suitability were established. DQ assessment is supported by risk assessment, with DQ being measured both from the perspective of the Data user as well as the perspective of Data production.

Overall Allianz's approach towards Data Governance is again well founded in theoretical Data Governance theories such as Weber et al (2008). It also borrows elements from the CDQM by having a dedicated DQM strategy as well as by developing a system for DQM (Otto et al, 2007). Its use of Data Quality metrics can also be seen in research contributions such as Hüner et al (2011).

5.1.5.2 Sparkassen Finanzgruppe

This practice report in the finance and insurance industry is provided by two articles of Finanz Informatik. These refer to both a report from Sparkasse KölnBonn & Markgräflerland, as well as more generally with regards to the same project of an integrated data household (IDH) at the Sparkassen Finanzgruppe. This project aims at providing the Data household with integrated business rules, allowing Data Quality Managers and Data Managers the necessary improvements. The project also includes an assessment team, as well as responsibilities for Master Data and the organizational function. It implements both the new roles of Data Quality Manager as well as the implementation of DQM processes. Although the second report states that around 100 rules are used in 160.000 rule cycles in the first half of 2019, however, no description is given of post or past project value analysis.

Overall, this project has some theoretical foundation, with the focus on a DQM system with business rules and an integrated DQM processes having some foundation in the CDQM framework for Data Quality Management (Otto et al, 2007). In this, it refers to the Corporate Data Quality architecture layer. Its use of dedicated roles for Data Quality Management, such as the Data Quality Manager, can also be seen as being based in Data Governance research contributions such as Weber et al (2008).

5.1.5.3 Credit Suisse

Both Helfert & Hermann (2002), although anonymously, and Winter et al (2003) describe the implementation of a meta-data based DQ system for the Data Warehouse at Credit Suisse. Due to strong similarities of the two systems as well as similar descriptions of the company the assumption was made that Helfert & Hermann (2002) also conducted their research at Credit Suisse. Business and as well as technical Data rules form an important element of the implemented system. These rules are based upon the experience values, with the DQ system being applied daily at the deployment layer of the Data Warehouse. Besides this system for DQM, an organizational process for DQ was also specified at Credit Suisse. It is reactive, either being triggered by data problems or by certain rules in the system being violated. This process also assigns the responsibilities and necessary actions, with the two main responsible roles being the IT DQ responsible and the respective DQ responsible from the affected department. The IT DQ responsible however, is initially responsible for the decision if the incident is of a technical nature and can be solved by the IT department or, if it is required, the cooperation of the business department.

Overall Credit Suisse's approach towards DQM is founded in theory, with its focus on a specific system for DQM referring to the governance and execution layers of the Corporate Data Quality (CDQ) architecture, which form one of the three areas of Corporate Data Quality Management according to Otto et al (2007). Credit Suisse also incorporates elements from the area of Corporate Data Quality organization of the CDQM in its development of processes and procedures for DQM as well as definition of roles and responsibilities. This is however limited to the governance layer of CDQ organization, with no mention of steps on the execution area such as e.g. training of employees.

5.1.5.4 Barmenia

The practice report by ACT IT Consulting & Services GmbH (2011) describes DQM at Barmenia insurance company. At Barmenia, DQM has the goals of improving DQ, establish a cross-functional Data Quality process and foster a feeling of responsibility for Data. In order to achieve this, the tool InfoZoom was used among others to visualize customer Data and its quality problems. Besides that, Barmenia implemented logic trees to identify root causes, conducted information events, and used a valuation method for DQM measures. The project for DQM was situated at the corporate IT function but was mostly conducted by external consultants. It was deemed successful although the practice report did not present a clear post project value analysis.

Overall, this project can be seen as having some basis in the CDQM framework, with the use of a dedicated DQM tool on the CDQ architecture layer as well as of a cross-functional process on its organizational layer (Otto et al, 2007). Its use of information events also refers to CDQ organization.

5.1.6 Utilities

Organization	References	Type
Water Inc.	Weber et al (2008)	Research contribution
Pfalzwerke Netzgesellschaft mbH	Everding (2010)	Practice report
SWB Energie & Wasser	Meyer (2012)	Practice report
Stadtwerke München	Trumpetter (2015)	Practice report
Stadtwerke Winsen (Luhe) GmbH	Meyer (2015)	Practice report

Table 29 Utility companies

5.1.6.1 Water Inc.

The large utilities company of Water Inc. is described in the context of its organization of accountabilities for DQM, as Data Quality is necessary due to regulatory requirements (Weber et al, 2008). DM at Water Inc. took the form of a team of Data Stewards, which found that a more detailed DG organization was needed in order to switch from what they perceived as a reactive approach to a proactive approach. Therefore, a Data governance organization was devised, consisting of a Data Governance council, a Data Custodian, a Data Steward and User groups consisting of Data stakeholders.

Water Inc's Data Quality Management approach yet again is based mostly on Data Governance as established by Weber et al (2008). Water Inc. establishes many of the described roles, such as the Data Governance council and Data Stewards.

5.1.6.2 Pfalzwerke Netzgesellschaft mbH

In the case of Pfalzwerke Netzgesellschaft GmbH, usage of a DQ tool, InfoZoom, is described. It is used to analyze DQ problems from their main SAP system (Everding, 2010). This analysis is conducted by assessing the SAP Data entries based on certain logical rules. The identified DQ errors can then be exported, allowing the Data owing departments to correct the errors.

This practice report regarding DQM at the Pfalzwerke Netzgesellschaft mbH provides only very limited Information with regards to the organization of DQM. It only focuses on a particular system for Data Quality Management. Therefore, the only theoretical foundation for its approach in assuring Data Quality is the Corporate Data Quality (CDQ) architecture of Otto et al (2007). While no information is provided with regards to the governance layer of the CDQ architecture, such as the development of a common Information object model, the use of InfoZoom can be related to the execution layer. As described by Otto et a (2007), Pfalzwerke Netzgesellschaft mbH operates a system for DQ analysis and cleaning, Info-Zoom.

5.1.6.3 SWB Energie & Wasser

Another utility company, SWB Energie und Wasser, is also mentioned with regards to the tool of InfoZoom (Meyer, 2012). Here, it is utilized for the analysis Master Data in the customer service and billing department. At SWB Energie und Wasser, InfoZoom is however also used for plausibility analyses, as well as for the correction of large amounts of false entries.

DQM at SWB Energie und Wasser however is only loosely founded in established theory of Data Quality Management. Its focus on a system for DQ can again be seen as most closely being founded in the Corporate Data Quality Architecture (Otto et al, 2007). As in the case of Pfalzwerke Netzgesellschaft mbH no information is provided regarding the governance layer of CDQ, with the operation of a system for DQ analysis and -improvement being based in the execution area of the CDQ architecture (Otto et al, 2007).

5.1.6.4 Stadtwerke München

Stadtwerke München, after failing with single and uncoordinated DQM initiatives combined them into a single project and institutionalized them (Trumpetter, 2015). This project was conducted on the layers of DQM organization, DQM tools and DQM organization, with a value analysis also being conducted. The DQM organizational area referred to the DQM team being incorporated into the shared services division, with dedicated DQ tasks and responsibilities such as Data Owner as well as dedicated committees such as the DQ steering committee. DQM processes at Stadtwerke München on the other hand referred to the design of a specific DQ process based upon business requirements. These processes included, (1) accepting the DQ requirement and defining the DQ, (2) measuring the DQ requirement, (3) analysis of DQ problems, (4) improvement of DQ problems and (5) monitoring and maintenance of achieved DQ improvements (Trumpetter, 2015). Ultimately, DQ tools refers to specific systems for DQM such as a DQ cockpit utilized to visualize DQ problems and automated error reports. Stadtwerke München also conducted extensive value analysis, confirming the efficiency and sustainability of their unified DQM initiative.

The DQM initiative at Stadtwerke München was also awarded second place in the CC DQM Best Practice awards and is heavily influenced by the CDQM framework. It is organized alongside the layers of the CDQM, but its processes for DQM also share a link to Total Data

Quality Management, them consisting of definition, measurement, analysis and improvement (Wang, 1998). Therefore, it is solidly founded from a research perspective.

5.1.6.5 Stadtwerke Winsen (Luhe) GmbH

The last contribution towards this issue is made with regards to Stadtwerke Winsen (Luhe) GmbH (Meyer, 2015). This initiative started with an analysis of the existing Data, with master Data being identified as an issue. In order to address this a DQM process was established by the Shared Services department. This process had to be implemented, including the assignment of business-oriented Data Owners. In order to address the technical necessities of DQM, a dedicated DQM system was purchased and implemented, the already mentioned InfoZoom. This allows the Data Owners to assess DQ without the need for in-depth technical knowledge. The DQM initiative and InfoZoom initiative was deemed successful, with various improvements in DQ leading to operational benefits such as customer satisfaction.

Overall, this initiative at least borrows elements of existing DQM concepts such as e.g. TDQM with analysis being the first step of the Data Quality Management process (Wang, 1998). It is also based on the Corporate Data Quality Management framework as described by Otto et al (2007), with the definition of DQM process as well as of responsibilities from the CDQ organization area. The CDQ architecture area can also be seen in part, with the adoption and use of a dedicated system for DQM.

5.1.7 Chemicals

Organization	**References**	**Type**
Ciba Inc.	Weber (2009)	Research contribution
Lanxess	Österle & Otto (2016)	Research contribution

Table 30 Chemical companies

5.1.7.1 Ciba Inc.

Another industry being referred to in the analyzed references is the chemical industry, such as Ciba Inc (Weber, 2009). Ciba Inc. started an operational excellence initiative due to regulatory requirements, with one goal being the implementation of a unified IT infrastructure and an SAP system. Other goals were the consolidation of Master Data, the implementation of DG roles and responsibilities, the formalization of related processes and the documentation of core Data objects. A Master Data organization was adopted, consisting of (1) a stewardship organization, (2) a Data maintenance organization and (3) a Data ownership organization. Therefore, the Data standards department was newly implemented in order to management Master Data and the Master Data organization. No prior value analysis of the actions related to these goals was conducted, as well as no qualitative post-analysis. While overall improvements were described, the actual quality of the master Data was not improved.

Overall, the DQM approach of Ciba Inc's is based in Data Governance, with distinctive roles for DQM being assigned (Weber et al, 2008). It also incorporates many elements of Corporate Data Quality Management according to Otto et al (2007), with the areas of CDQ strategy, CDQ organization and CDQ system all being present. Ciba Inc. develops both a strategic IT plan and ensures regulatory compliance on the governance area, however still lacking the description of goal communication on the execution layer of CDQ strategy. This is similar in the CDQ organization area, where also only processes on the governance layer are described. On the CDQ architecture area very little information is provided with regards to Ciba Inc., but its goals of consolidating MD as well as documenting core Data objects can be seen as referring to this area.

5.1.7.2 Lanxess

A more detailed description of DQM in the chemical industry at Lanxess is provided by Österle & Otto (2016). Lanxess conducted an IT consolidation process from 2004 to 2011, which concluded in one master Data system, two ERP systems and one global reporting system. Besides this, a global department to support Master Data Management activities was also implemented, as well as a Data Governance organization devised. The MDM department among other tasks, was responsible for training of employees in the business departments for MD usage and its processes. Besides this, another organization was created for the Data Owners, however even after these measures, Master Data could still be changed by to many people. Therefore, the REMIX projects for Business Intelligence and Reporting was conducted, with utilized pilot studies and resulted in the identification of the need for a unified view on Master Data as well as for a single source of truth in its regard. Besides this, it was also found that in-memory computing systems needed to be used for BI and Reporting due to large amount of Data.

Overall, Data Quality Management at Lanxess is based both in theories of Data Governance and of Corporate Data Quality Management (Weber et al, 2008; Otto et al, 2007). Roles are assigned for DQM, but also DQM organizational elements implemented. The provision of training for employees with regards to MD can also be seen in being founded in the Corporate Data Quality organization execution area of Otto et al (2007).

5.1.8 Other industries

Organization	References	Type	Industry
Bosch	Österle & Otto (2016)	Research contribution	Electric engineering
Festo	Österle & Otto (2016)	Research contribution	Automotive supplier
Hilti	Österle & Otto (2016)	Research contribution	Construction
Shell	Österle & Otto (2016)	Research contribution	Energy
Hartje	Eggheads GmbH (2020)	Practice report	Manufacturing

Table 31 Other industries

5.1.8.1 Bosch

Bosch is described by Österle & Otto (2016) as having a focus on Data architecture with regards to DQM and as having implemented a corporate policy for master Data management. This policy differentiated between organizational and technical Master Data Management activities, with organizational MDM activities being responsibility of the Data Owners. The IT department is responsible for the technical ones. Bosch also utilizes four corporate functions, those being (1) Data Governance, (2) Data Provisioning, (3) Data Usage and (4) Data Management concepts & projects. Bosch is also described as utilizing a singular Master Data architecture, however with differences for the different kinds of master Data.

Overall, Bosch's approach is most closely again based in Corporate Data Quality Management (Otto et al, 2007). Responsibilities for Data are assigned in the area of CDQ organization, while the utilization of a singular architecture for Master Data can be linked to the CDQ architecture area of CDQM.

5.1.8.2 Festo

A company in the automotive supplier industry, Festo, is mentioned by Österle & Otto (2016), with it needing more efficient Master Data Management. Its current project in this regards aims at reorganizing the related administrative tools, provide libraries and reorganize the Master Data organization. In this project a staging area for the PLM system was implemented as well as four success factors for MDM identified, those being (1) the creation of a shared consciousness for MDM, (2) the design of business processes, (3) Tool support and (4) centralized Product Data administration. Besides this, it was also learned that harmonized business processes support Product Data Quality and that an integrated Product Lifecycle Management system is needed, as well as that responsibility for MDM has to be shared.

Overall, Festo's approach towards DQM is very specific and undetailed with regards to the organizational and strategic aspects of DQM. That is described, is mostly founded in the system layer of the CDQM (Otto et al, 2007).

5.1.8.3 Hilti

Another company described by Österle & Otto (2016) in the construction industry, Hilti, again focused on customer master Data. It is described as initially having no corporate-wide DQM and as having three risks in customer Data management. A project was conducted in order to design an implemented a Customer Data Quality Tool, which was financed by the market reach department. The tool was based upon business rules devised by subject matter experts. Overall goals of the project also included the creation of a shared consciousness for MD, as well as of transparency and DQ monitoring. In this context, Hilti differentiated between proactive and reactive actions, with the proactive actions being the definition of roles and responsibilities for customer master Data and to correct customer Data creation errors. Reactive actions on the other hand include DQ monitoring, Data maintenance and a tool for the cleaning of mass Data. Project success was found to be dependent on top management support as well as already discussed factors such as business rule definition.

Overall, Hilti provides another example of a very system focused approach towards Data Quality Management. In this, it is founded in the CDQ architecture area according to Otto et al (2007).

5.1.8.4 Shell

Shell, a company in the energy industry, focused its approach towards DQM on Product Lifecycle Data. The Product Lifecycle Management in this case was reengineered via Six Sigma, with six different projects under unified leadership. These projects included (1) improving system support for PLM, (2) definition of global and local business rules, (3) developing of a tool for automated filling out of entry fields, (4) implementation of a SAP Master Data solution, (5) implementation of a workflow management system for production and (6) reduction of manual actions in product creation. The value gained from these projects was assessed beforehand via a Design Measure Analyze Improve (DMAIC) approach.

Overall, the case of Shell provides very little information with regards to the organization of the described DQM measures. The described projects can be seen as gain referring to the system layer of Data Quality Management (Österle & Otto, 2016). Its utilization of the DMAIC process for pre-project value evaluation however provides a link to Total Data

Quality Management. In TDQM, Data Quality Management follows similar steps, except what it refers to the first stage as definition of the information product (Wang, 1998).

5.1.8.5 Hartje

A manufacturing company, Hartje, is describe by eggheads GmbH (2020) as lacking centralized Data as well as the ability to measure Data Quality. For this, a new system was adopted, eggheads Suite, which allows for centralized product Data as well as Data export and import. It also contains a DQM module.

Overall, the case of Hartje again highlights a focus on DQM systems and the architecture layer of the CDQM framework (Otto el, 2007). It is also another example of only a very brief and technical description being given in a practice report.

5.2 Data Quality Management differences by organization types

Besides the issue of the use of DQM in different industries, the differences between organizations types are also of consideration and e.g. described by Sautter et al (2018). They differentiate between enterprises, research institutions and cities. Enterprises, unlike research institutions according to contain a singular objective in the provision of goods or services and are more homogeneous. On the other hand, department and organizational elements of research institutes are more independent of each other than in cities. No mention of DQM was found with regards to cities in the reviewed references and practice reports, with e.g. the identified utilities companies being enterprises owned by cities but not strictly a part of them. Since however enterprises make out the vast majority of the companies mentioned in the DQM context, further differentiation is needed in their regard.

5.2.1 Enterprises

As mentioned above, enterprises form the vast majority of the organizations referred to in DQM research contributions and practical sources, making them somewhat of a baseline for the practical expression of DQM. Therefore, differentiation will be made between large national/multinational companies (NC) such as e.g. Allianz and smaller and more regional companies (RC) such as Water Inc. and Stadtwerke München (Österle & Otto, 2016; Weber et al, 2008; Trumpetter, 2015). Large national/multinational companies' expression of DQM puts a heavier focus on harmonization and centralization of DQM activities and processes, this being necessary due to the greater complexity and heterogeneity of their IT and

Application architecture through growth as well as M&A activities. They also put heavier emphasis on pre and/or post DQM project value analysis. Overall, they employ more of a Data Governance approach, assigning roles and responsibilities for Data classes and types but not without providing details with regards to the actual processes of DQM.

Regional companies on the other hand, are described with a heavier focus on specific DQM tools and systems such as InfoZoom, with the only company with a more extensive and detailed DQM initiative being Stadtwerke München, a much larger company than the other three. This heavier focus leads to them being very much described on the system layer of the Corporate Data Quality framework, with them either not focusing on CDQ strategy and organization or the reports being too limited in this regard. The report of Stadtwerke Munich hints at the second, with it explicitly also referring to the strategy and organization layers.

5.2.2 Research institutions

Besides enterprises however, there is only one research contribution with regards to DQM expression in research institutions, provided by Leadbetter et al (2020). Leadbetter et al (2020) provide a much more detailed description of the implementation of DQM at the Irish Marine Institute. Their approach had to comply to regulatory requirements and focused on the implementation of DG roles such as Data Owner and Data Protection officer. Although only two sources with regards to this organizational type were identified, both this lack of prevalence in the practical and research sources as well as the findings of Houston et al (2018) hints at a lack of focus on DQM in research institutions. They found among other things a significant lack of organization of DQM in clinical research sites (Houston et al, 2018). Another possibility is a lack of communication and publication in this regards. This is reinforced by the fact that both sources are significantly newer than the average research contribution

6 Challenges and Future research topics

Ultimately, the unsolved issues and as well as challenges with regards to Data Quality Management have to be addressed and summarized. These challenges can be differentiated in terms of Data-related and management-related ones. Besides this, the possible future research topics can be derived from these challenges.

6.1 Data Challenges and Future research topics

The first type of DQM challenges are the ones being related to Data itself. In this context, three Data-related challenges were found.

Challenge	References
Unstructured Data	Shankaranarayanan & Blake (2017), Madnick et al (2009)
Big Data	Shankaranarayanan & Blake, (2017), Cai & Zhu (2015), Frehe et al (2016), Jaya et al (2017)
Semantic Integration	Österle & Otto (2016)

Table 32 Data related challenges

6.1.1 Unstructured Data

One Data-related challenge for DQM is unstructured Data. Shankaranarayanan & Blake (2017) describe Data increases from social media and increases in transactional Data as a challenge with regards to Data Quality. This Data is described as much less structured than the Master Data which is the subject of many common research contributions, such as the ones made by the CC DQM. This challenge for DQM is also mentioned by Madnick et al (2009), according to which additional research is needed in order to develop techniques and methods to manage the quality of this type of Data. Such research might e.g. be conducted with Social Media companies and incorporate anonymous case studies and/or action research in order to identify how these companies assure the quality of this unstructured Data processes by them.

6.1.2 Big Data

Big Data is a major challenge of DQM according to Shankaranarayanan & Blake, (2017), with Cai & Zhu (2015) also specifically mentioning Big Data as a challenge for DQM. This is due to its inherent factors such as volume, velocity, veracity, and variety (Shankaranarayanan & Blake, 2017; Cai & Zhu). Another factor in Big Data, according to Frehe et al (2016), is its value. Therefore, all five of these factors pose a possible challenge for DQM.

(1) Volume: Big Data's enormous volume possess a challenge for DQM, with it posing a challenge for the DQM systems to process it in an acceptable timeframe.

(2) Velocity: The velocity of Big Data possess another challenge for DQM in that the growth of Data volume is both dynamic and fast. This might pose a challenge for

DQM systems as well as DQM processes attuned to a more stable and constant flow of master Data.

(3) Veracity: Yet another challenge of DQM with regards to Big Data is that of its veracity. Big Data's quality according to Frehe et al (2016) is often already questionable in terms of uncertainty and credibility, decreasing DQ already before time-related deterioration.

(4) Variety: The variety of Big Data also possess a challenge, due to many DQM approaches in the literature and practice reports making usage of business rules. These business rules might not be applicable to Big Data due to it being much more varied than the Master Data they were designed for.

(5) Value: Perhaps, the last challenge with regards to Big Data is how to derive value from DQM action in its regard. The value derived from DQM initiatives targeting master Data such as customer Data is much more tangible and measurable, than the value derived from Big Data whose value even at high DQ is muss less tangible.

These factors are less of an issue in many current approaches in achieving Data Quality due to their focus on master Data, which is both more stable and constant in volume than transactional Data in the Big Data context. An early research contribution towards the issue of Data Quality and Big Data is provided by Merino et al (2016) by their 3-As Model, but nonetheless this challenge warrants future research by e.g. further empirical validation of the 3-As Model.

6.1.3 Semantic integration

Another Data-related DQM challenge according to Österle & Otto (2016) is that of semantic integration. This refers to the need for corporate Data to be used in a unified manner, for which DQM needs to integrate it. This can be done by combining them into a single system. It can also be done by having diverse system which are linked via interfaces. However, the advantages and disadvantages of each approach towards sematic integration have not been a research topic in the reviewed literature. This may be addressed through some form of comparative analysis e.g. by conducting a survey with Data Stewards with companies utilizing the different approaches.

6.2 Managerial Challenges and Future research topics

Challenges	References
DQM investment justification	Österle & Otto (2016); Francisco et al (2017), Baghi (2017)
Organizational complexity	Österle & Otto (2016)
Cloud computing	Al-Ruithe et al (2019)
DQM resource allocation	Leadbetter et al (2020), Österle & Otto (2016)
Compliance	Österle & Otto (2016), Al-Ruithe et al (2019)
Lack of business involvement	Niemi (2017)

Table 33 Managerial challenges

6.2.1 DQM investment justification

On the managerial side a common challenge identified in the practical expression of DQM is the need to justify DQM activities by analyzing their value. This is described in several case studies of Österle & Otto (2016). While many of the discussed organizations conduct pre-and/or post project value analysis, other organizations conducted DQM measures without clear and quantifiable performance indicators. This hints at a possible challenge of clearly defining the value gained from DQM in some instances, which might hinder DQM initiatives in enterprises with reluctant leadership or sparse financial resources. Similarly, Francisco et al (2017) also conclude their research with the suggestion of a development of a method to measure financial returns through DQ measures. This is also mentioned by Baghi (2017), who also recommends the design of method to value DQM actions. Falge (2014) already includes a phase for value analysis in their method for strategy development, but also hints at a need for further research in this regard. Therefore, the development of a specific method for the valuation of DQM actions may be an important future research topic.

6.2.2 Organizational complexity

Another identified challenge for DQM is that of a very heterogeneous and spread out IT landscape and its corresponding complexity, as described in cases such as the one of Shell by Österle & Otto (2016). Both through growth by expansion to new countries or regions as well as through M&A activities, an enterprise's IT and ERP landscape becomes less centralized and unified over time, leading to decreases in Data Quality. It is also more generally described as challenge by Österle & Otto (2016) with regards to centralized Data architectures. Centralized DQM and MDQM functions, a unified view on Master Data, Data Governance elements as well as dedicated DQM system might be able to counteract DQM challenges through organizational complexity and assure Data Quality. New acquisitions however and/or rapid enterprise growth might always pose a challenge with regards to the assuring all time high Data Quality. However, the Data architecture still has to be flexible enough to handle new requirements and handle both internal and external Data objects. This challenge may be addressed by e.g. a quantitative survey conducted with regards to the efficiency and effectives of DQM in different companies, with factors such as number of employees, number of divisions, IT architecture approach etc. also being collected.

6.2.3 Cloud computing

Another managerial challenge according to Al-Ruithe et al (2019) is the implementation of Data Governance in the cloud context. A differentiation is made between technical, legal, and business issues in this regard:

(1) **Technical challenges:** Technical challenges regarding the implementation of Data Governance in Cloud architectures include security, privacy, availability, performance, Data classification and Data migration.
(2) **Legal challenges:** According to Al-Ruithe et al (2019) legal challenges refer to many legal factors influencing the implementation of Data Governance, such as the need for compliance with differing regulations. Besides this, the complexity of the legal contracts also poses a challenge for Data owners and customers in the cloud context.
(3) **Business challenges:** These challenges for Data Governance in the cloud context refer to organizational factors which influence the DG implantation, such as top management support and organizational size. Due to this, there is no single approach towards the implementing Data Governance in cloud settings.

This challenge for Data Governance and DQM may be addressed by e.g. conducting an exploratory case study in companies that incorporate both Data Governance and DQM functions as well utilize cloud services for their technical DQM operations.

6.2.4 Resource allocation for DQM

Another managerial challenges found in the reviewed references is the allocation of sufficient resources to DQM initiatives and activities. Leadbetter et al (2020) describes the main challenge of their DQM implementation as the resourcing of the Data Owners and Stewards. These key DQM actors had to be freed from their every-day task in order to support the DQM initiative. Similarly, Österle & Otto (2016) describe one of the main challenges for

Shell's DQM initiative in insufficient allocation of personal resources to the project with regards to the tool development. Besides this they also state that the projects funding was limited in comparison to similar projects, leading to a need to focus on key functionalities. This warrants further research into the amount of resources allocated to DQM initiatives, via e.g. a qualitative survey on companies which conducted DQM projects.

6.2.5 Compliance

The compliance to regulation poses another challenge for DQM. Österle & Otto (2016) describe one of the Challenges of Shell's DQM initiative in the need to convince its internal audit and compliance team of its compliance to regulations and government requirements. It also more generally describes the need for Data security as challenge for Data Management. Similarly, Al-Ruithe et al (2019) also describe the need for compliance as a challenge with regards to the implementation of Data Governance. This challenge for DQM may be subject to future research via e.g. conducting surveys with Data Stewards whose companies have adopted DQM and DG functions and operate in highly regulated areas of operations.

6.2.6 Lack of Business Involvement

Niemi (2017) describes a lack of business involvement as another managerial challenge for DQM. According to the results of the conducted literature analysis business involvement serves to establish the context in which the Data is to be used, however Data is still often perceived as just an IT problem. In order for the Data to have the necessary contextual quality DQM requires involvement from the business departments. This challenge of business involvement may lead to new research regarding the opinions and predispositions of Data owners, e.g. via a survey regarding their perception of responsibility towards Data maintenance.

7 Conclusion

Data Quality Management as a research field is diverse, incorporating both more technical approaches as well as more managerial and business-related approaches. This contribution aims at providing an overview of the exiting DQM research, its concepts, topics, and research methodologies. Based on this sample of DQM research contributions, the assumption can be made that the TDQM project and the CC DQM project constitute a substantial degree of Data Quality Management research. Many non-associated contributions nonetheless refer to one or more of their research contributions. These different projects also somewhat showcase the development of DQM, starting with it as a sequential process of Data Quality improvement according to the TDQM and developing into a constant corporate objective, which has to be incorporated and requires consideration from different areas of the enterprise, according to the CDQM.

This developing nature can also be seen in the utilized methodologies, with a heavy focus on Design Science and Case study research with regards to DQM. Frameworks and other forms of practical artifacts are developed, and case studies conducted in their regard. This hints at the necessity for academic contributions to be practical and possible to implement in enterprises being dependent on high quality Data. Big Data and increasing amounts of transactional Data are developing challenges in this regard, requiring a re-shift from the focus on corporate and Master Data. Ultimately, DQM requires a combination of managerial and technical skills within the organizations, as well as a solid governance framework in its regard. Due to the increases in Data volume as well as an increasing focus on Data analytics for process optimizations the overall importance of Data Quality Management is likely to only increase in the future.

References

- ACT IT-Consulting & Services GmbH. (2011). Datenqualitätsmanagement bei der Barmenia Versicherung. unn | UNITED NEWS NETWORK GmbH. https://www.pressebox.de/pressemitteilung/act-gruppe/Datenqualitaetsmanagement-bei-der-Barmenia-Versicherung/boxid/415759 (visited last 30.08.2020)
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2019). A systematic literature review of Data governance and cloud Data governance. *Personal and Ubiquitous Computing, 23(5-6)*, 839–859. https://doi.org/10.1007/s00779-017-1104-3
- Bai, X. (2012). A Mathematical Framework for Data Quality Management in Enterprise Systems. *INFORMS Journal on Computing*, *24*(4), 648–664. https://doi.org/10.1287/ijoc.1110.0475
- Balzert, Helmut; Schröder, Marion; Schäfer, Christian; Motte, Petra (2015): Wissenschaftliches Arbeiten. Ethik, Inhalt & Form wiss. Arbeiten, Handwerkszeug, Quellen, Projektmanagement, Präsentation. 2. Auflage, 4. Nachdruck, Sonderdruck für die Studierenden der Leibniz-FH. Dortmund: W3L-Verlag (Soft Skills).
- Bargh, M. S., Mbgong, F., van Dijk, J., & Choenni, S. (2015). A FRAMEWORK FOR DYNAMIC Data QUALITY MANAGEMENT. *In EC 2015 Proceedings: Proceedings of the IADIS International Conference e-Commerce and Digital Marketing* (pp. 134–142).). http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=114760634&site=ehost-live
- Batini, C., Barone, D., Mastrella, M., Maurino, A., & Ruffini, C. (2007). A Framework And A Methodology For Data Quality Assessment And Monitoring. In *ICIQ* (pp. 333–346).
- Batini, Carlo; Cappiello, Cinzia; Francalanci, Chiara; Maurino, Andrea (2009): Methodologies for Data Quality assessment and improvement. In: *ACM Comput. Surv.* 41 (3), S. 1–52.
- Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly, 11(3)*, 369. https://doi.org/10.2307/248684
- Bernd Heinrich, Diana Hristova, Mathias Klier, Alexander Schiller, & Michael Szubartowicz (2018). Requirements for Data Quality Metrics. *Journal of Data and Information Quality, 9(2)*, 1–32. https://doi.org/10.1145/3148238

- Caballero, I., Caro, A., Calero, C., & Piattini, M. (2008). IQM3: Information Quality Management Maturity Model. *J. UCS, 14(22)*, 3658–3685.
- Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. Data *Science Journal, 14(0)*, 2. https://doi.org/10.5334/dsj-2015-002
- Coughlan, P., & Coghlan, D. (2002). Action research for operations management. *International Journal of Operations & Production Management, 22(2)*, 220–240. https://doi.org/10.1108/01443570210417515
- Dalmolen, S., Moonen, H., & van Hillegersberg, J. (2015). Industry-wide Inter-organizational Systems and Data Quality: Exploratory findings of the use of GS1 standards in the Dutch retail market. In AMCIS. https://aisel.aisnet.org/amcis2015/e-Biz/GeneralPresentations/32
- Dresch, A., Lacerda, D. P., & Miguel, P. A. C. (2015). A distinctive analysis of case study, action research and design science research. *Revista Brasileira De Gestão De Negócios, 17(56)*, 1116.
- Edelen, A., & Ingwersen, W. W. (2018). The creation, management, and use of Data Quality Information for life cycle assessment. *The International Journal of Life Cycle Assessment, 23(4)*, 759–772. https://doi.org/10.1007/s11367-017-1348-1
- eggheads GmbH. (2020). Hartje: Jetzt mit zentralisierten Produktdaten Vollgas geben. unn | UNITED NEWS NETWORK GmbH. https://www.pressebox.de/pressemitteilung/eggheads-gmbh/Jetzt-mit-zentralisierten-Produktdaten-Vollgas-geben/boxid/1001973 (visited last 30.08.2020)
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62(1)*, 107–115.
- Everding, S. (2010, August 24). Pfalzwerke Netzgesellschaft mbH | SAP Daten analysieren, Datenqualität optimieren – intuitive Analysen ohne Suchvorgaben. HumanIT Software GmbH. https://www.infozoom.com/sap-daten-datenqualitaet-pfalzwerke-netzgesellschaft/
- Falge, C. (2014). Methode zur Strategieentwicklung für unternehmensweites Datenqualitätsmanagement in globalen Konzernen [@St. Gallen, Univ., Diss., 2014]. http://www1.unisg.ch/www/edis.nsf/wwwDisplayIdentifier/4349
- Falge, C., Otto, B., & Österle, H. (2012). Data Quality Requirements of Collaborative Business Processes. In R. H. Sprague (Ed.), 2012 45th Hawaii International Conference

on System Science: (HICSS) ; USA, 4 - 7 Jan. 2012 (pp. 4316–4325). IEEE. https://doi.org/10.1109/HICSS.2012.8

- Falge, C., Otto, B., & Österle, H. (2013). Towards a Strategy Design Method for Corporate Data Quality Management. In Wirtschaftsinformatik (p. 51).
- finanz Informatik. (2020, August 22). Datenqualität ist messbar. https://www.f-i.de/News/ITmagazin/Archiv/2018/Integrierter-Datenhaushalt/Titelthema/Datenquali-taet-ist-messbar (visited last 30.08.2020)
- finanz Informatik. (2020, August 22). Neues vom Integrierten Datenhaushalt. https://www.f-i.de/News/ITmagazin/Archiv/2019/FI-Connect-2019-VorausDenkenMachen/Loesungen-Praxis/Neues-vom-Integrierten-Datenhaushalt (visited last 30.08.2020)
- Francisco, M. M. C., Alves-Souza, S. N., Campos, E. G. L., & Souza, L. S. de. (2017). Total Data Quality Management and Total Information Quality Management Applied to Customer Relationship Management. In Proceedings of the 9th International Conference on Information Management and Engineering (pp. 40–45). https://doi.org/10.1145/3149572.3149575
- Frehe, V., Adelmayer, T., & Teuteberg, F. (2016). Eine Balanced Scorecard für das systematische Datenqualitätsmanagement im Kontext von Big Data. In Multikonferenz Wirtschaftsinformatik 2016.
- Gable, G. G. (1994). Integrating case study and survey research methods: an example in Information systems. *European Journal of Information Systems*, *3*(2), 112–126.
- Glowalla, P., & Sunyaev, A. (2013). Process-Driven Data Quality Management Through Integration of Data Quality into Existing Process Models. *Business & Information Systems Engineering*, *5*(6), 433–448. https://doi.org/10.1007/s12599-013-0297-x
- Glowalla, P., & Sunyaev, A. (2014). Process-driven Data Quality Management: A Critical Review on the Application of Process Modeling Languages. *Journal of Data and Information Quality (JDIQ)*, *5*(1-2), 1–30. https://doi.org/10.1145/2629568
- Grimmer, U., & Hinrichs, H. (2001). A Methodological Approach to Data Quality Management Supported by Data Mining. *Proceedings of the Sixth International Conference on Information Quality*.

- Heinrich, B., & Klier, M. (2006). Ein Optimierungsansatz für ein fortlaufendes Datenqualitätsmanagement und seine praktische Anwendung bei Kundenkampagnen. *Journal of Business Economics(76(6))*, 559–587.
- Houston, L., Probst, Y., Yu, P., & Martin, A. (2018). Exploring Data Quality Management within Clinical Trials. *Applied Clinical Informatics, 9(1)*, 72–81. https://doi.org/10.1055/s-0037-1621702
- Hüner, K. (2011) (I). Method for Specifying Business-oriented Data Quality Metrics (BE HSG / CC CDQ / 28) [, St. Gallen]. www.alexandria.unisg.ch. (I) https://www.alexandria.unisg.ch/84077/
- Hüner, K., Ofner, M., & Otto, B. (2009). Towards a maturity model for corporate Data Quality Management. In *Proceedings of the 2009 ACM symposium on Applied Computing*.
- Jansen, H. (2010). The logic of qualitative survey research and its position in the field of social research methods. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research. Vol. 11. No. 2.*
- Jaya, I., Sidi, F., Affendey, L., Jabar, M., & ISHAK, I. (2019). SYSTEMATIC REVIEW OF Data QUALITY RESEARCH. *Journal of Theoretical and Applied Information Technology, 97*, 3043.
- Jaya, M. Izham, Sidi, F., ISHAK, I., Suriani Affendey, L., & JABAR, M. A. (2017). A REVIEW OF Data QUALITY RESEARCH IN ACHIEVING HIGH Data QUALITY WITHIN ORGANIZATION. *Journal of Theoretical & Applied Information Technology, 95(12)*.
- Kreis, L. (2017). Datenqualität als kritischer Erfolgsfaktor bei Datenmigrationen. ZHAW Zürcher Hochschule für Angewandte Wissenschaften.
- Kwon, O., Lee, N., & Shin, B. (2014). Data Quality Management, Data usage experience and acquisition intention of big Data analytics. *International Journal of Information Management, 34(3)*, 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002
- Laranjeiro, N., Soydemir, S. N., & Bernardino, J. (2015). A Survey on Data Quality: Classifying Poor Data. In 2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC).
- Leadbetter, A., Carr, R., Flynn, S., Meaney, W., Moran, S., Bogan, Y., Brophy, L., Lyons, K., Stokes, D., & Thomas, R. (2020). Implementation of a Data Management

Quality Management Framework at the Marine Institute, Ireland. *Earth Science Informatics, 13(2)*, 509–521. https://doi.org/10.1007/s12145-019-00432-w

- Liaw, S.-T., Pearce, C., Liyanage, H., Liaw, G. S. S., & Lusignan, S. de (2014). An integrated organization-wide Data Quality Management and Information governance framework: Theoretical underpinnings. *Informatics in Primary Care, 21(4)*, 199–206. https://doi.org/10.14236/jhi.v21i4.87
- Lucas, A. (2010). Corporate Data Quality Management: From theory to practice. *In 5th Iberian Conference on Information Systems and Technologies.*
- Lucas, A. (2010). Towards corporate Data Quality Management. *Portuguese Journal of Management Studies, XV(2)*, 173–196.
- M. Hüner, K., Schierning, A., Otto, B., & Österle, H. (2011). Product Data Quality in supply chains: The case of Beiersdorf. *Electronic Markets, 21(2)*, 141–154. https://doi.org/10.1007/s12525-011-0059-x
- Madnick, S. E., Wang, R. Y., Lee, Y. W., & Zhu, H. (2009). Overview and framework for Data and Information Quality research. *Journal of Data and Information Quality (JDIQ), 1(1)*, 1–22.
- Madnick, S., Wang, R., & Xian, X. (2003). The design and implementation of a corporate householding knowledge processor to improve Data Quality. *Journal of Management Information Systems, 20(3)*, 41–70.
- March, S. T., & Storey, V. C. (2008). Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research. *MIS Quarterly, 32(4)*, 725. https://doi.org/10.2307/25148869
- Merino, J., Caballero, I., Rivas, B., Serrano, M., & Piattini, M. (2016). A Data Quality in Use model for Big Data. *Future Generation Computer Systems, 63*, 123–130. https://doi.org/10.1016/j.future.2015.11.024
- Meyer, G. (2012, August 7). SWB Energie und Wasser | Datenqualitätsmanagement SAP Daten: Datenkontrolle im Fachbereich. HumanIT Software GmbH. https://www.infozoom.com/datenqualitaetsmanagement-sap-daten-swb-energie-und-wasser/ (visited last 30.08.2020)
- Meyer, G. (2015, April 23). Stadtwerke Winsen (Luhe) GmbH | Datenqualitätsmanagement: Wertschöpfung durch saubere Daten. HumanIT Software GmbH. https://www.infozoom.com/datenqualitaetsmanagement-stadtwerke-winsen-luhe-gmbh/

- Ofner, M. H., Hüner, K. M., & Otto, B. (2009). Dealing with complexity: a method to adapt and implement a maturity model for corporate Data Quality Management. *AMCIS 2009 Proceedings*, 491.
- Ofner, M., Otto, B., & Österle, H. (2013). A Maturity Model for Enterprise Data Quality Management. *Enterprise Modelling and Information Systems Architectures, Vol 8, No 2* (2013)
- Österle, H., & Otto, B. (2016). Corporate Data Quality: Voraussetzung erfolgreicher Geschäftsmodelle. Springer. http://www.doabooks.org/doab?func=fulltext&rid=19124
- Otto, B. (2011). Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers. *Communications of the Association for Information Systems, 29*. https://doi.org/10.17705/1CAIS.02903
- Otto, B., & Ebner, V. (2010). Measuring Master Data Quality: Findings from an Expert Survey. In Multikonferenz Wirtschaftsinformatik.
- Otto, B., & Hinderer, H. (2009). Datenqualitätsmanagement im Lieferanten-Controlling. *Controlling & Management, 53(1)*, 21–29.
- Otto, B., & Hüner, K. (2009). Functional Reference Architecture for Corporate Master Data Management (BE HSG / CC CDQ / 21) [, St. Gallen]. www.alexandria.unisg.ch. https://www.alexandria.unisg.ch/214250/
- Otto, B., Hüner, K. M., & Österle, H. (2012). Toward a functional reference model for master Data Quality Management. *Information Systems and E-Business Management*, *10*(3), 395–425. https://doi.org/10.1007/s10257-011-0178-0
- Otto, B., Wende, K., Schmidt, A., & Osl, P. (2007). Towards a framework for corporate Data Quality Management. In 18th Australasian Conference in Information Systems.
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data Quality assessment. *Communications of the ACM, 45(4)*, 211–218.
- Ryu, K. S., Park, J. S., & Park, J. H. (2006). A Data Quality Management Maturity Model. *ETRI Journal, 28(2)*, 191–204.
- Sautter, J., Litauer, R., Fischer, R., Klages, T., Wuchner, A., Müller, E., Schaj, G., Dobrokhotova, E., Drews, P., & Riess, S. (2018). Beyond Data Quality: Data Excellence Challenges from an Enterprise, Research and City Perspective. Data, 245–252.

- Schäffer, T., & Stelzer, D. (2017). Assessing Tools for Coordinating Quality of Master Data in Inter-organizational Product Information Sharing. In Wirtschaftsinformatik 2017.
- Schäffer, T., Leyh, C., & Beckmann, H. ALADDIN–Vorschlag eines Analyse- und Berechnungsmodells zur Investitionsbewertung für ein unternehmensweites Datenqualitätsmanagement.: Data driven X–Turning Data into Value. In Multikonferenz Wirtschaftsinformatik (MKWI) (Vol. 2018).
- Schmidt, A., Hüner, K., & Grewe, A. (2010). Fallstudie Deutsche Telekom AG-Einheitliche Datenarchitektur als Grundlage für unternehmensweites Datenqualitätsmanagement. RIS.
- Shamala, P., Ahmad, R., Zolait, A., & Sedek, M. (2017). Integrating Information Quality dimensions into Information security risk management (ISRM). *Journal of Information Security and Applications*, *36*, 1–10. https://doi.org/10.1016/j.jisa.2017.07.004
- Shankaranarayan, G., Ziad, M., & Wang, R. Y. (2003). Managing Data Quality in dynamic decision environments: An Information product approach. *Journal of Database Management (JDM)*(14(4),), 14-32.
- Shankaranarayanan, G., & Blake, R. (2017). From Content to Context: The Evolution and Growth of Data Quality Research. *Journal of Data and Information Quality (JDIQ)*, 8(2), 1–28. https://doi.org/10.1145/2996198
- Shankaranarayanan, G., & Cai, Y. (2006). Supporting Data Quality Management in decision-making. *Decision Support Systems, 42(1)*, 302–317. https://doi.org/10.1016/j.dss.2004.12.006
- Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data Quality in context. *Communications of the ACM*, *40*(5), 103–110.
- Trumpetter Joerg. (2015). Etablierung eines nachhaltigen Datenqualitätsmanagements bei den Stadtwerken München.
- Wand, Y., & Wang, R. Y. (1996). Anchoring Data Quality dimensions in ontological foundations. *Communications of the ACM, 39(11),* 86–95.
- Wang, R. Y. (1998). A product perspective on total Data Quality Management. *Communications of the ACM, 41(2)*, 58–65. https://doi.org/10.1145/269012.269022
- Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems, 12(4)*, 5–33.

- Wang, R. Y., Storey, V. C., & Firth, C. P. (1995). A framework for analysis of Data Quality research. *IEEE Transactions on Knowledge and Data Engineering, 7(4)*, 623–640.
- Weber, K. (2009). Data Governance-Referenzmodell: Organisatorische Gestaltung des unternehmensweiten Datenqualitätsmanagements [Diss. Nr. 3731 Wirtschaftswiss. St. Gallen, S.l.]. Schweizerische Nationalbibliothek.
- Weber, K., Cheong, L. K., Otto, B., & Chang, V. (2008). Organizing Accountabilities for Data Quality Management - A Data Governance Case Study. In B. Dinter (Ed.), *GI-Edition Proceedings: Vol. 138, Synergien durch Integration und Informationslogistik: DW2008 ; 27./28.10.2008 in St. Gallen, Switzerland* (pp. 347–359). Ges. für Informatik. https://www.alexandria.unisg.ch/67463/
- Weber, K., Otto, B., & Österle, H. (2009). Data GOVERNANCE: ORGANISATIONSKONZEPT FÜR DAS KONZERNWEITE DATENQUALITÄTSMANAGEMENT: Doctoral dissertation, Difo-Druck GmbH.(I)
- Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All---A Contingency Approach to Data Governance. *Journal of Data and Information Quality (JDIQ), 1(1)*, 1–27. (II)
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii.
- Weidema, B. P., & Wesnaes, M. S. (1996). Data Quality Management for life cycle inventories—an example of using Data Quality indicators. *Journal of Cleaner Production (4(3-4))*, 167-174.
- Wende, K. (2007). A Model for Data Governance – Organizing Accountabilities for Data Quality Management. *ACIS 2007 Proceedings*. https://aisel.aisnet.org/acis2007/80
- Westin, S., & Sein, M. (2015). The Design and Emergence of a Data/Information Quality System. *Scandinavian Journal of Information Systems, 27(1)*. https://aisel.aisnet.org/sjis/vol27/iss1/1
- Winter, M., Herrmann, C., & Helfert, M. (Ed.). (2003). Data Warehouse Management: Das St. Galler Konzept zur ganzheitlichen Gestaltung der Informationslogistik.
- Würthele, V. (2003). Datenqualitätsmetrik für Informationsprozesse: [Datenqualitätsmanagement mittels ganzheitlicher Messung der Datenqualität] [Diss. Nr. 15242 Wiss. ETH Zürich, S.l.]. Schweizerische Nationalbibliothek.

- Zhang, R., Sadiq, S., & Indulska, M. (2019). Discovering Data Quality Problems. *Business & Information Systems Engineering, 61(5)*, 575–593. https://aisel.aisnet.org/bise/vol61/iss5/3
- Zhu, H., Madnick, S. E., Lee, Y. W., & Wang, R. Y. (2014). Data and Information Quality Research: Its Evolution and Future. In *Computing Handbook Set.*

Appendices

Appendix A: References excluded due criterion 3

Authors	Titles	Year	Source
Lee, Y. W., Pipino, L., Funk, J. D., & Wang, R. Y.	Journey to Data Quality	2006	Book
Otto, B., Hüner, K., & Österle, H.	A Cybernetic View on Data Quality Management	2010	*AMCIS*
Röthlin, M.	Management of Data Quality in enterprise resource planning systems	2010	Book
Otto, B.	Quality Management of Corporate Data Assets.	2011	Quality management for IT services: Perspectives on business and process performance (pp.193-209)
Schumacher & Weiß	Prozess- und Data Governance als strategischer Ansatz zur Verbesserung der Prozess- und Datenqualität in Unternehmen	2011	HMD Praxis der Wirtschaftsinformatik, 48(3)
Fan, W., & Geerts, F.	Foundations of Data Quality Management	2012	Synthesis Lectures on Data Management, 4(5
Otto, B.	On the Evolution of Data Governance in Firms: The Case of Johnson & Johnson Consumer Products North America.	2013	Handbook of Data Quality: Research and practice (pp. 93-118)
Hildebrand, K., Gebauer, M., Hinrichs, H., Mielke, M.	Daten- und Informationsqualität: Auf dem Weg zur Information Excellence	2015	Book
S. Juddoo	Overview of Data Quality challenges in the context of Big Data.	2015	2015 International Conference on Computing, Communication and Security (ICCCS).
Weigel, N.	Datenqualitätsmanagement - Steigerung der Datenqualität mit Methode	2015	Daten- und Informationsqualität: Auf dem Weg zur Information Excellence (pp. 69–86)

Batini, C., & Scannapieco, M.	Data and Information Quality: Dimensions, principles and techniques. Data-centric systems and applications	2016	Book
Carretero, A. G., Caballero, I., & Piattini, M.	MAMD: Towards a Data Improvement Model Based on ISO 8000-6X and ISO/IEC 33000	2016	International Conference on Software Process Improvement and Capability Determination
Otto, B., & Legner, C.	Datenqualitätsmanagement für den Industriebetrieb: Best Practices und Implikationen der Digitalisierung	2016	Controlling: Zeitschrift für erfolgsorientierte Unternehmenssteuerung, 28(10)
Schäffer, T., & Leyh, C.	Master Data Quality in the Era of Digitization - Toward Inter-organizational Master Data Quality in Value Networks: A Problem Identification	2017	International Conference on Enterprise Resource Planning Systems
Bicevska, Z., Bicevskis, J., & Oditis, I.	Models of Data Quality.	2018	Lecture Notes in Business Information Processing: Vol. 311
Cruz, S. T., Gomez, A. F., Lopez Sevillano, A., & Lozano-Garzon, C.	Audit to the Data Quality Management Process in a Small Organization Based on NTC-ISO 19011	2018	2018 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI)

Appendix B: References excluded due to criterion 4

Authors	Titles	Year	Source
Tayi & Ballou	Examining Data Quality	1998	Communications of the ACM, 41(2), 54–57.
Loshin	Issues and Opportunities in Data Quality Management Coordination	2004	DM Review, 14(4), 14–16
Schmid	The main steps to Data Quality	2004	Industrial Conference on Data Mining.
Otto	Compliance und Data Governance im Stammdaten-Management	2007	IIR Stammdaten-Management Forum https://www.alexandria.unisg.ch/213100/
Otto	Strategischer Erfolgsfaktor Stammdatenqualität	2007	BearingPoint Fachtagung: Nutzenorientiertes Stammdaten- Management im SAP Umfeld - Wieviel Harmonisierung ist wirklich nötig? https://www.alexandria.unisg.ch/213145/
Niemi	Designing a Data Governance Framework	2011	*Proceedings of the IRIS Conference* (Vol. 14).
Otto et al	Stammdatenmanagement: Datenqualität für Geschäftsprozesse	2011	*HMD Praxis der Wirtschaftsinformatik*, *48*(3), 5–16.
Vancauwenbergh	Data Quality Management	2019	*Scientometrics Recent Advances.* IntechOpen. https://doi.org/10.5772/intechopen.86819

Appendix C: Survey research contributions

Reference	Year	Titles	Goal	Focus group	Sample size	Analysis
Wang, R. Y., & Strong, D. M.	1996	Beyond Accuracy: What Data Quality Means to Data Consumers	Develop a framework for capturing DQ aspects	Data users	1st stage: 137 2nd stage: 1500, 355 responses	Factor analysis
Otto, B., & Ebner, V.	2010	Measuring Master Data Quality: Findings from an Expert Survey	Survey of DQ assessment in companies	Employees related to DQM	300, 41 responses	-
Kwon, O., Lee, N., & Shin, B.	2014	Data quality management, data usage experience and acquisition intention of big data analytics	Influence factors on Big-Data analytics adoption	Companies	939, 306 responses	Structural equation modeling
Kreis, L.	2017	Datenqualität als kritischer Erfolgsfaktor bei Datenmigrationen	Success factors for Data Migration	People that participated in migration projects	53, 36 responses	Only descriptive
Shamala, P., Ahmad, R., Zolait, A., & Sedek, M.	2017	Integrating information quality dimensions into information security risk management (ISRM)	Identify relevant DQ dimensions for Information Security Risk Management	Practitioners from organizations with ISO compliance	201, 150 responses	Partial Least Square
Houston, L., Probst, Y., Yu, P., & Martin, A.	2018	Exploring Data Quality Management within Clinical Trials	Development of standard Data Quality monitoring procedures to ensure Data integrity	Clinical research sites	142, 20 responses	Only descriptive

Appendix D: Identified references (1/4)

Authors	Year	Title	Source	Topic	Method
Wang et al	1995	A framework for analysis of data quality research	IEEE transactions on knowledge and data engineering	DQ/DQM research	Literature review
Weidema & Wesnaes	1996	Data quality management for life cycle inventories—an example of using data quality indicators.	Journal of cleaner production	DQ assessment	Unclear
Wang & Strong	1996	Beyond Accuracy: What Data Quality Means to Data Consumers	Journal of Management Information Systems	DQ dimensions	Survey
Wand & Wang	1996	Anchoring data quality dimensions in ontological foundations	Communications of the ACM	DQ dimensions	Unclear
Strong et al	1997	Data quality in context	Communications of the ACM	DQ dimensions	Case study
Wang	1998	A product perspective on total data quality management	Communications of the ACM	TDQM	Unclear
Pipino et al	2002	Data quality assessment	Communications of the ACM	DQ assessment	Unclear
Würthele	2003	Datenqualitätsmetrik für Informationsprozesse	University paper	PDDQM	Case study
Winter et al	2003	Data Warehouse Management: Das St. Galler Konzept zur ganzheitlichen Gestaltung der	University paper	DQM system	Unclear
Shankaranarayan et al	2003	Managing data quality in dynamic decision environments: An information product approach	Journal of Database Management (JDM)	IMAP	Unclear
Madnick et al	2003	The design and implementation of a corporate householding knowledge processor to improve	Journal of management	DQM system	Unclear
Shankaranarayan & Cai	2006	Supporting data quality management in decision-making	Decision Support Systems	IMAP	Unclear
Ryu et al	2006	A Data Quality Management Maturity Model	ETRI Journal	DQM maturity	Survey
Heinrich & Klier	2006	Ein Optimierungsansatz für ein fortlaufendes Datenqualitätsmanagement und seine praktische Anwendung bei Kundenkampagnen	Journal of Business Economics	DQM efficiency	Case study
Wende	2007	A Model for Data Governance – Organising Accountabilities for Data Quality Management	ACIS 2007 Proceedings	DQM & Data Governance	Unclear
Otto et al	2007	Towards a framework for corporate data quality management	18th Australasian Conference in Information Systems	CDQM	Design Science
Batini et al	2007	A Framework And A Methodology For Data Quality Assessment And Monitoring	ICIQ	DQ assessment	Unclear
Caballero et al	2008	IQM3: Information Quality Management Maturity Model	J. UCS	DQM maturity	Unclear
Weber et al	2008	Organising Accountabilities for Data Quality Management - A Data Governance Case Study	GI-Edition Proceedings: Vol. 138, Synergien durch Integration und Informationslogistik: DW2008	DQM & Data Governance	Case study
Weber et al	2009	One Size Does Not Fit All---A Contingency Approach to Data Governance	Journal of Data and Information Quality (JDIQ)	DQM & Data Governance	Action research
Weber et al	2009	DATA GOVERNANCE: ORGANISATIONSKONZEPT FÜR DAS KONZERNWEITE	University paper	DQM & Data Governance	Design Science
Weber	2009	Data Governance-Referenzmodell: Organisatorische Gestaltung des unternehmensweiten Datenqualitätsmanagements	University paper	DQM & Data Governance	Unclear
Otto et al	2009	Functional Reference Architecture for Corporate Master Data Management	University paper	MDQM	Design Science

Appendix E: Identified references (2/4)

Otto & Hinderer	2009	Datenqualitätsmanagement im Lieferanten-Controlling	Controlling & Management	CDQM	Design Science
Ofner et al	2009	Dealing with complexity: a method to adapt and implement a maturity model for corporate data quality management	AMCIS 2009 Proceedings	DQM maturity	Design Science
Madnick et al	2009	Overview and framework for data and information quality research	Journal of Data and Information Quality (JDIQ)	DQ/DQM research	Unclear
Hüner et al	2009	Towards a maturity model for corporate data quality management	Proceedings of the 2009 ACM symposium on Applied Computing	DQM maturity	Design Science
Batini et al	2009	Methodologies for data quality assessment and improvement	ACM Comput. Surv. (ACM Computing Surveys)	DQ assessment	Comparative analysis
Schmidt et al	2010	Fallstudie Deutsche Telekom AG-Einheitliche Datenarchitektur als Grundlage für unternehmensweites Datenqualitätsmanagement	University paper	DQM system	Case study
Otto & Ebner	2010	Measuring Master Data Quality: Findings from an Expert Survey	Multikonferenz Wirtschaftsinformatik	DQ assessment	Survey
Lucas	2010	Corporate data quality management: From theory to practice	5th Iberian Conference on Information Systems and Technologies	DQM frameworks	Case study
Lucas	2010	Towards corporate data quality management	Portuguese Journal of Management Studies	DQM frameworks	Case study
Otto	2011	Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers	CAIS (Communications of the Association for Information Systems)	DQM & Data Governance	Case study
Hüner et al	2011	Product data quality in supply chains: the case of Beiersdorf	Electronic Markets	DQ metrics	Case study
Hüner	2011	Method for Specifying Business-oriented Data Quality Metrics	University paper	DQ metrics	Design Science
Grimmer & Hinrichs	2011	A Methodological Approach to Data Quality Management Supported by Data Mining	Proceedings of the Sixth International Conference on Information Quality	DQM frameworks	Design Science
Otto et al	2012	Toward a functional reference model for master data quality management	Inf Syst E-Bus Manage (Information Systems and e-Business Management)	MDQM	Design Science

Appendix F: Identified references (3/4)

Authors	Year	Title	Source	Topic	Method
Falge et al	2012	Data Quality Requirements of Collaborative Business Processes	45th Hawaii International Conference	DQ effects	Qualitative content analysis
Bai	2012	A Mathematical Framework for Data Quality Management in Enterprise Systems	Informs Journal of Computing	DQ metrics	Case study
Ofner et al	2013	A Maturity Model for Enterprise Data Quality Management	Enterprise Modelling and Information Systems Architectures	DQM maturity	Design Science
Glowalla & Sunyaev	2013	Process-Driven Data Quality Management Through Integration of Data Quality into Existing Process Models	Business & Information Systems Engineering	PDDQM	Literature review
Falge et al	2013	Towards a Strategy Design Method for Corporate Data Quality Management	Wirtschaftsinformatik	DQM strategy	Design Science
Liaw et al	2014	An integrated organisation-wide data quality management and information governance framework: theoretical underpinnings	Informatics in primary care	DQM & Data Governance	Literature review
Kwon et al	2014	Data quality management, data usage experience and acquisition intention of big data analytics	International Journal of Information Management	DQ effects	Survey
Glowalla & Sunyaev	2014	Process-driven data quality management: A critical review on the application of process modeling languages	Journal of Data and Information Quality (JDIQ)	PDDQM	Literature review
Falge	2014	Methode zur Strategieentwicklung für unternehmensweites Datenqualitätsmanagement in globalen Konzernen	University paper	DQM strategy	Design Science
Westin & Sein	2015	The Design and Emergence of a Data/Information Quality System	Scandinavian Journal of	DQM system	Action research
Laranjeiro et al	2015	A Survey on Data Quality: Classifying Poor Data	2015 IEEE 21st Pacific Rim 2015		
Dalmolen et al	2015	Industry-wide Inter-organizational Systems and Data Quality: Exploratory findings of the use of GS1 standards in the Dutch retail market	AMCIS 2015 Proceedings	DQ effects	Case study
Bargh et al	2015	A FRAMEWORK FOR DYNAMIC DATA QUALITY MANAGEMENT	International Conference on Information Systems Post-Implementation & Change Management	DQM frameworks	Design Science
Österle & Otto	2016	Corporate Data Quality: Voraussetzung erfolgreicher Geschäftsmodelle	Monography	CDQM	Unclear
Merino et al	2016	A Data Quality in Use model for Big Data	Future Generation Computer Systems	DQ assessment	Design Science
Frehe et al	2016	Eine Balanced Scorecard für das systematische Datenqualitätsmanagement im Kontext von Big Data	Multikonferenz Wirtschaftsinformatik 2016	DQM frameworks	Design Science

Appendix G: Identified references (4/4)

Shankaranarayan & Blake	2017	From content to context: The evolution and growth of data quality research.	Journal of Data and Information Quality (JDIQ)	DQ/DQM research	Laten Semantic Analysis
Shamala et al	2017	Integrating information quality dimensions into information security risk management (ISRM)	Journal of Information	DQ dimensions	Survey
Schäffer & Stelzer	2017	Assessing Tools for Coordinating Quality of Master Data in Inter-organizational Product Information Sharing	Wirtschaftsinform atik	DQM system	Case study
Nurminen	2017	EFFECTIVE CORPORATE DATA QUALITY MANAGEMENT: Systematic Literature Review	University paper	DQ/DQM research	Literature review
Kreis	2017	Datenqualität als kritischer Erfolgsfaktor bei Datenmigrationen	University paper	DQ effects	Survey
Jaya et al	2017	A REVIEW OF DATA QUALITY RESEARCH IN ACHIEVING HIGH DATA QUALITY WITHIN ORGANIZATION	Journal of Theoretical & Applied Information Technology	DQ/DQM research	Literature review
Francisco et al	2017	Total Data Quality Management and Total Information Quality Management Applied to Costumer Relationship Management	Proceedings of the 9th International Conference on Information Management and Engineering	TDQM	Comparative analysis
Baghi	2017	Capability reference model for establishing data quality controlling	University paper	DQM efficiency	Design Science
Schäffer et al	2018	ALADDIN–Vorschlag eines Analyse-und Berechnungsmodells zur Investitionsbewertung	Multikonferenz Wirtschaftsinform	DQM efficiency	Design Science
Sautter et al	2018	Beyond Data Quality: Data Excellence Challenges from an Enterprise, Research and City Perspective	DATA	DQ/DQM research	Literature review
Houston et al	2018	Exploring Data Quality Management within Clinical Trials	Applied Clinical Informatics	DQ/DQM research	Survey
Heinrich et al	2018	Requirements for Data Quality Metrics	Journal of Data and Information	DQ metrics	Unclear
Edelen & Ingwersen	2018	The creation, management, and use of data quality information for life cycle assessment	The international journal of life cycle	DQ assessment	Unclear
Zhang et al	2019	Discovering Data Quality Problems	Business & Information Systems Engineering	DQ assessment	Design Science
Jaya et al	2019	SYSTEMATIC REVIEW OF DATA QUALITY RESEARCH	Journal of Theoretical and Applied Information Technology	DQ/DQM research	Literature review
Al-Ruithe et al	2019	A systematic literature review of data governance and cloud data governance	Personal and Ubiquitous	DQ/DQM research	Literature review
Leadbetter et al	2020	Implementation of a Data Management Quality Management Framework at the Marine Institute,	Earth Science Informatics	DQM frameworks	Unclear

Erklärung

Hiermit versichere ich, dass diese Abschlussarbeit von mir persönlich verfasst ist und dass ich keinerlei fremde Hilfe in Anspruch genommen habe. Ebenso versichere ich, dass diese Arbeit oder Teile daraus weder von mir selbst noch von anderen als Leistungsnachweise andernorts eingereicht wurden. Wörtliche oder sinngemäße Übernahmen aus anderen Schriften und Veröffentlichungen in gedruckter oder elektronischer Form sind gekennzeichnet. Sämtliche Sekundärliteratur und sonstige Quellen sind nachgewiesen und in der Bibliographie aufgeführt. Das Gleiche gilt für graphische Darstellungen und Bilder sowie für alle Internet-Quellen.

München, den 6. October 2020

S. Rodler

Unterschrift