

A Data Quality Management Framework to Support Delivery and Consultancy of CRM Platforms

Renee Albrecht, Sietse Overbeek and Inge van de Weerd

*Department of Information and Computing Sciences, Faculty of Science, Utrecht University,
Princetonplein 5, 3584 CC Utrecht, The Netherlands*

Keywords: Consultancy, CRM, Data Quality Management, Delivery, Design Science.

Abstract: CRM platforms heavily depend on high-quality data, where poor-quality data can negatively influence its adoption. Additionally, these platforms are increasingly interconnected and complex to meet growing needs of customers. Hence, delivery and consultancy of CRM platforms becomes highly complex. In this study, we propose a CRM data quality management framework that supports CRM delivery and consultancy firms to improve data quality management practices within their projects. The framework should also improve data quality within CRM solutions for their clients. We extract best practices for CRM data quality management by means of a literature study on data quality definition and measurement, data quality challenges, and data quality management methods. In a case study at an IT consultancy company, we investigate how CRM delivery and consultancy projects can benefit from the incorporation of data quality management practices. The design of the framework is validated by means of confirmatory focus groups and a questionnaire. The results translate into a framework that provides a high-level overview of data quality management practices incorporated in CRM delivery and consultancy projects. It includes the following components: Client profiling, project definition, preparation, migration/integration, data quality definition, assessment, and improvement.

1 INTRODUCTION

Contemporary enterprises heavily depend on data, where data is seen as a strategic asset (Nagle et al., 2020). The International Data Corporation (IDC) states that during 2020 the amount of data that has been created, captured, and replicated across the world is more than 59 zettabytes, and over the next five years the world will create more than three times the amount of data compared to the previous five years (IDC, 2020). With an increasing volume, velocity, and variety of data, the issues arising from errors in data and the organizational impact of these issues are amplified (Laney, 2018). Laney states that “poor data quality can have grave consequences, from strategic decisions that can lead to the death of a business to operation decisions that can lead to the death of individuals”, and: “[...] 40 percent of all failed business initiatives are a result of poor data quality” (Laney, 2018, pp. 246–247). Poor Data Quality (DQ) is one of the greatest challenges facing contemporary enterprises (Davenport and Harris, 2017). Simultaneously, enterprises struggle to address their data issues, while high-quality data is rather the exception than the

rule (Nagle et al., 2020). By managing DQ, unwanted consequences can be prevented, and valuable insights can be discovered in regards to interactions with customers.

An important operational aspect of enterprises that relies upon its utilisation of high-quality data is Customer Relationship Management (CRM). Poor DQ and integration can negatively influence the adoption of CRM (Cruz-Jesus et al., 2019). Additionally, survey data collected from about 300 organisations for the *State of CRM Data Management* report shows that 44% of its respondents estimate a loss in revenue as a result of poor quality CRM data, which ranges from 5%-20% of total revenue (Hanson, 2020). CRM is defined as “the core business strategy that aims to create and maintain profitable relationships with customers, by designing and delivering superior value proposition” (Buttle and Maklan, 2019, p. 21). It is enabled by information technology in the form of CRM platforms, at present often provided by IT consultancy firms as CRM (cloud) solutions. Those firms offer Delivery & Consultancy (D&C) of the CRM platform to a variety of customers. Those customers vary and grow in their needs, business processes, and goals.

Consequently, the CRM platforms are increasingly interconnected and complex, resulting in a continuous need for more study in the area of the management of CRM platform development, implementation, and marketing (Stone et al., 2017).

The objective of this research is to provide a solution for the need for data of adequate quality in increasingly complex CRM platforms. The aim is to improve data quality within CRM platforms by designing a framework that supports data quality management in order to keep the quality of CRM data on an adequate level. This translates to the following research question: *How can a data quality management framework be designed to support CRM platform delivery and consultancy?*

The following section introduces the research approach. Subsequently, we present a CRM DQ management framework in section 3, followed by its validation in section 4. The results are discussed in section 5 and, finally, we present our conclusions after which possibilities for further research are elaborated on in section 6.

2 RESEARCH APPROACH

In this study, we adopt a design science approach. Our aim is to investigate and design a framework for data quality management in CRM platform D&C. In line with Wieringa's Design Science Framework (Wieringa, 2014), we follow an iterative set of problem-solving tasks according to the structure of the so-called design cycle.

The first task is *Problem Investigation*, where the design of the CRM DQ Management Framework (CRM-DQMF) is prepared by conducting exploratory research to understand the problem. To increase the robustness of the results of this research, methodological triangulation is adopted (Kaplan and Maxwell, 2005). We carried out a literature review to identify, evaluate, and integrate findings of relevant high-quality studies that address the research problem. We determined the relevance of the literature by scope, objectives, methods, and conclusions (Budgen and Brereton, 2006). To ensure a complete review, additional search methods that are used are forward and backward searching (Levy and Ellis, 2006). We also included grey literature in order to get insights on the state of the art concerning DQ management in CRM platforms in practice.

The second part of our problem investigation was carried out in the form of a single embedded case study at an IT consultancy firm (Yin, 2003). The goal of this case study was to investigate the defined prob-

lem within its context. Its results are triangulated by including exploratory expert interviews with a total of 14 experts, as well as a documentation analysis including 15 relevant documents mainly existing of company documentation. The participating experts and researched documentation can be found in table 1 and table 2 respectively.

Table 1: Participants Expert Interviews.

ID	Role	Years of Experience
P1	Consultant	4
P2	Senior Analyst	2
P3	Senior Director	14
P4	Group Manager	15
P5	Senior Consultant	13
P6	Senior Consultant	5
P7	Senior Analyst	4
P8	Manager	9
P9	Consultant	4
P10	Consultant	6
P11	Analyst	2
P12	Senior Consultant	7
P13	Director	14
P14	Senior Consultant	7

Table 2: Researched documentation.

ID	Purpose	Year
D1	CRM platform insights	2020
D2	Data migration guidelines	2020
D3	CRM (on-premises) to cloud migration guidelines	2020
D4	CRM platform adoption guidelines	2020
D5	Common Data Model	2020
D6	Report on the solution design #1 for a utility client	2020
D7	Report on the solution design #2 for a utility client	2020
D8	Data flows explanation within a solution for a utility client	2020
D9	Functional design datamart for a utility client	2020
D10	Education in CRM platform functionalities	2021
D11	Global policy in data management	2021
D12	Dedicated system for data quality checks in multiple systems	2011
D13	Data management roadmap proposal	2021
D14	Data quality analysis strategy at a product client	2019
D15	Design principles in a CRM project of an NGO client	2021

The second task, *Treatment Design*, specifies requirements for the CRM-DQMF, which are extracted from insights of the triangulated collected data. Based

on those requirements, the design of the CRM-DQMF is established. The results are presented by means of a Process-Deliverable Diagram (PDD) consisting of two integrated diagrams, namely a process diagram including all activities, and a deliverable diagram including the deliverables that result from the activities (van de Weerd and Brinkkemper, 2008; van de Weerd, 2009). We used an assembly-based method engineering approach that facilitates situational analysis and design methods. Since situational factors play a key role when managing DQ, e.g. the CRM platform, the industry, and data processes, this is deemed as a suitable approach for the design of the CRM-DQMF.

In the last task, *Treatment Validation*, we validate the initial design of the CRM-DQMF by means of expert opinions extracted using confirmatory focus groups with a total of 6 experts (Tremblay et al., 2010). Subsequently, we conducted an interactive questionnaire among the same group of experts to effectively validate the first design of the CRM-DQMF (Robson and McCartan, 2016).

This study has several contributions. First, the investigation of CRM and its D&C provides insights in what CRM D&C projects entail and which practices have to be taken into consideration. Second, the understanding of which types of data are utilised in contemporary CRM platforms serves as input on how a definition of CRM DQ can be established, as well as how CRM DQ can be measured. Third, the investigation of known challenges regarding DQ within CRM platforms and the potential solutions contributes to the development of criteria that DQ management within CRM D&C should adhere to. Lastly, the review of existing DQ management methods based on extracted requirements for their applicability to CRM D&C projects results in a list of criteria for DQ management in CRM D&C projects.

3 RESULTS

From the expert interviews as well as the documentation analysis as part of the case study it can be concluded that the level of DQ management in contemporary CRM D&C projects is lacking, as there is no mutual awareness of the importance of DQ management in clients or CRM D&C employees, nor are there best practices in place to perform DQ management integrated in CRM D&C projects. Experts that participated in interviews state that CRM D&C teams face the challenge of the existing client's context, with its own constraints and levels of expertise regarding DQ and DQ management. DQ management practices are not standardized within CRM D&C projects

at the IT company of the case study, and are applied and performed subjectively varying depending on the experience(s) of the expert. Additionally, experts stated that there is need for DQ management, as found challenges in CRM D&C projects are largely related to poor DQ. Experts mutually experienced that currently, assessment and improvement of DQ is not done proactively, while they agree that this would benefit the D&C project. Ideally, to provide a complete and accurate as possible solution and advice as a CRM D&C team to the client, DQ management needs to be taken into account in every CRM D&C project, from the start of the project.

The case study contributes to the findings of the literature review in two different ways: (1) By confirming literature findings on DQ management and CRM D&C projects, and (2) by providing new insights on the current level of DQ management in CRM D&C projects, the level of necessity for the topic, and the manner on which DQ management could be integrated in CRM D&C projects. From the findings in literature and the case study, the requirements for the CRM-DQMF as explained in the following section are extracted. This is followed by an elaboration on the individual components of the CRM-DQMF in section 3.2.

3.1 Requirements Specification

For the CRM-DQMF to assist CRM D&C teams effectively in the management of DQ, it should adhere to certain criteria that ensure the management of DQ in general, as well as the management of DQ in the context of CRM D&C projects. It includes *Modularity, DQ Management Plan, DQ Management Maturity Level, CRM D&C Client Context, Migration/Integration, Iteration, Business Impact Analysis, DQ Assessment, and DQ Improvement*.

3.1.1 Modularity

Customers expect a CRM platform to be tailored for their organisation specifically, with limited effort, and deployed promptly (Cricelli et al., 2019). Insights as found in the case study indicate customer-centric and agile approaches for CRM D&C projects. Hence, in order to make DQ management in CRM D&C projects succeed, the CRM-DQMF is required to be designed in a modular fashion. Modularity refers to the uniqueness of every client and CRM D&C project, and therefore the need for uniqueness in DQ management application. The CRM-DQMF should be applicable in varying situations serving different needs. To serve this requirement, the CRM-DQMF exists of different components that can be separated and/or

(re)combined when required. This way, the CRM D&C team can be either the executive or advising force of a component, or take a more passive role and omit the component of the CRM-DQMF, leaving the responsibility entirely to the client. The modular visualization is facilitated by making use of situational method engineering as proposed by van de Weerd and Brinkkemper (van de Weerd, 2009). By the introduction of an activity which produces a CRM D&C project specific DQ management plan as explained in section 3.1.2, the remainder of the utilisation of the CRM-DQMF is decided upon.

3.1.2 DQ Management Plan

The CRM-DQMF facilitates the establishment of a unique DQ management plan at the start of any project. It describes the roles and responsibilities of the client and the CRM D&C team with regards to the required DQ management practices for the CRM D&C project. The DQ management plan is established based on the matters that make a client and project unique. This includes a business case, as this formats the required DQ management practices. The business case comprises of the client's budget, the client's business goals of a CRM D&C project, and the scope of the project. The budget indicates to what extent the client will be able to pay for DQ management services. The business goals provide an indication of the need for DQ management. The scope of the project indicates which functionalities are required for reaching the business goals, thus provides for an indication of the extent to which DQ management is required. Additionally, this includes the impact of DQ on the business goals, and the current DQ management maturity level and goals of the client (see section 3.1.3).

3.1.3 DQ Management Maturity Level

Results of the experts interviews indicate that DQ, and therefore the need for its management within a CRM D&C project along with the role of the CRM D&C team, depends on the expertise of the client. P3 mentioned: "In the ideal case, organisations already have an authority in place that takes care of data quality matters. However, this varies per organisation and industry. The interference of us depends on the arrangements with the client". Additionally, documentation mentions various DQ Key Performance Indicators, such as the number of data elements with a definition. This indicates that there is need for the determination of the client's DQ management maturity level, which determines the extent to which DQ management will be applied and by whom. On the

one hand, the client might not have any knowledge on their own data, nor its quality, which might indicate that the data is not of sufficient quality for a CRM solution and the client does not have sufficient in-house DQ management expertise, meaning the expertise of the CRM D&C team is required. On the other hand, the client might already be in control of its data (and quality) across the organisation, which means there is no need for the CRM D&C team to conduct or advise on any DQ management practices. Therefore, the current DQ management maturity level of the client as well as the goal DQ maturity level of the client play relevant roles for the establishment of the DQ management plan. To determine the DQ management maturity level of the client, the maturity matrix by Spruit and Pietzka can be utilised (Spruit and Pietzka, 2015). The capabilities of this maturity matrix are confirmed by this research. Those capabilities read: *Assessment of DQ*; *Impact on Business*; *Root causes of poor DQ*; and *DQ Improvement*. For each capability, the client can be at another maturity level, reading from lowest to highest: *Initial*; *Repeatable*; *Defined Process*; *Managed & Measurable*; and *Optimized*.

3.1.4 CRM D&C Client Context

Taking upfront considerations into account is considered to be important for DQ management in CRM D&C projects (Batini et al., 2009). This is confirmed by the case study in terms of the definition of a client context. Documentation showed the establishment of a so-called blueprint of the project, which is defined to create the framework for the CRM solution based on budget, goal, and scope. Experts mentioned this blueprint, as well as practices for the reconstruction of business processes of the client. The client context comprises information for a reconstruction of the organisational environment in regards to the CRM solution, which includes business processes, data, data policies, and data standards. Conventionally, a client context is already established by the CRM D&C team and the client as general part of the CRM D&C project.

3.1.5 Migration/Integration

CRM D&C projects comprise the process towards a CRM solution, which typically includes the migration and/or integration of data. P4 explained that "Business requirements are defined in order to decide which data is required for the solution. From this, a data mapping is realized to load the data to the new CRM solution correctly". Migration and integration practices comprise the mapping of data of two different systems: a legacy system and the new CRM so-

lution (Ali et al., 2017). Therefore, the CRM-DQMF should contain guidance to integrate DQ management into migration and integration practices, including data mapping practices. The practices of the *Transform* phase of the ETL process which is required for data migration and integration practices as found in literature (Thalheim and Wang, 2013) are confirmed by experts and documentation of the case study. This includes the mapping of the data, as well as the assessment and improvement of DQ.

3.1.6 Iteration

In terms of the CRM D&C project, iteration needs to take place to successfully establish migration and integration, as the case study indicates that the migration and integration consist of continuous gathering and refinement of business requirements, business rules, and data mappings. In terms of DQ management, this requires iterations of DQ assessment and improvement practices (Batini et al., 2009; Cichy and Rass, 2019). Experts mention phrases such as “I think data quality should be measured frequently in any case”, and in documentation as well as by experts the data lifecycle is mentioned, which consists of the creation, management, and destruction of data. This indicates that DQ should be managed up until its destruction. Once the CRM solution is established, the CRM-DQMF should still provide for iterations, as the assessment and improvement phases of the CRM-DQMF need to be ongoing processes, which solely end in case the lifecycle of all concerned data within the scope of the CRM D&C solution ceases to exist.

3.1.7 Business Impact Analysis

The business impact of poor DQ needs to be analysed, as this defines the data elements that are critical for the client’s business goals and thus require DQ assessment and potentially improvement practices (Heinrich et al., 2018; Cichy and Rass, 2019; Batini et al., 2009). This is also referred to as a *top down* or *demand driven* approach. Due to the variety in CRM D&C projects and clients, the eleven business impacts of poor DQ as identified by Spruit and van der Linden are recommended to be included within the impact analysis, as they are found to be applicable for a variety of industries (Spruit and Linden, 2019). Furthermore, the case study indicates agile and customer-centric project approaches. By not exclusively including monetary impact, the business impacts support an agile project approach, as agile values an emphasis on the quality, the flexibility and the customer-centricity of services (Rosing and Gill, 2015) and cost efficiency is not at the centre of atten-

tion in an agile project approach (Gill and Henderson-Sellers, 2006). The business impacts include lost sales opportunities, customer service costs, customer dissatisfaction, lost revenue, operational deficiencies, delays in system/project deployment, regulatory compliance, poor decision making, lost business opportunities, employee moral, and system credibility. For each significant business impact, a metric has to be defined to calculate its value (Batini et al., 2007).

3.1.8 DQ Assessment

Within DQ management, DQ assessment is found to be a critical part and should take place frequently (Cichy and Rass, 2019; Batini et al., 2009). Therefore, the CRM-DQMF should include guidance in assessment practices. This includes the definition and measurement of DQ and reporting on potential DQ issues. DQ can have different definitions, which is highly context dependent (Otto et al., 2007; Wang, 1998). Each client and project is unique, meaning DQ is required to be defined for every project. A DQ definition is expressed in terms of DQ dimensions and DQ thresholds (Otto et al., 2007; Wang, 1998; Cichy and Rass, 2019). The unstructured, semi-structured, and structured data used by CRM need to be taken into account when defining the appropriate DQ definition (Zahay et al., 2012; Missi et al., 2005). For each DQ definition, DQ metrics need to be defined to be able to quantify DQ (Pipino et al., 2002; Cichy and Rass, 2019). Additionally, for DQ assessment and improvement to succeed, the assignment of data roles is required (Batini et al., 2009). From the case study can be concluded that, without someone responsible for the data, its quality will not be managed.

3.1.9 DQ Improvement

When data is found to be of insufficient quality, the CRM-DQMF should offer guidance in establishing an improvement strategy (Batini et al., 2009; Cichy and Rass, 2019). For optimal DQ, an organisation needs to be aware of different reasons for poor DQ and where they are existent within the organisation, hence the root causes of DQ issues need to be analysed (Spruit and Pietzka, 2015; Batini et al., 2009; Cichy and Rass, 2019). The whereabouts of the weak spots should be known, as well as the reason(s) for the existence of weak spots, serving as input for the establishment of an improvement strategy. Direct and indirect costs of DQ are compared by making use of a cost evaluation to support a decision making process for its development of an improvement strategy (Spruit and Pietzka, 2015). These costs include the costs of the business impacts of the DQ issues, as well

as the costs of potential improvement practices.

3.2 Activities & Deliverables

The entire CRM-DQMF comprises seven phases, as can be seen in figure 1.

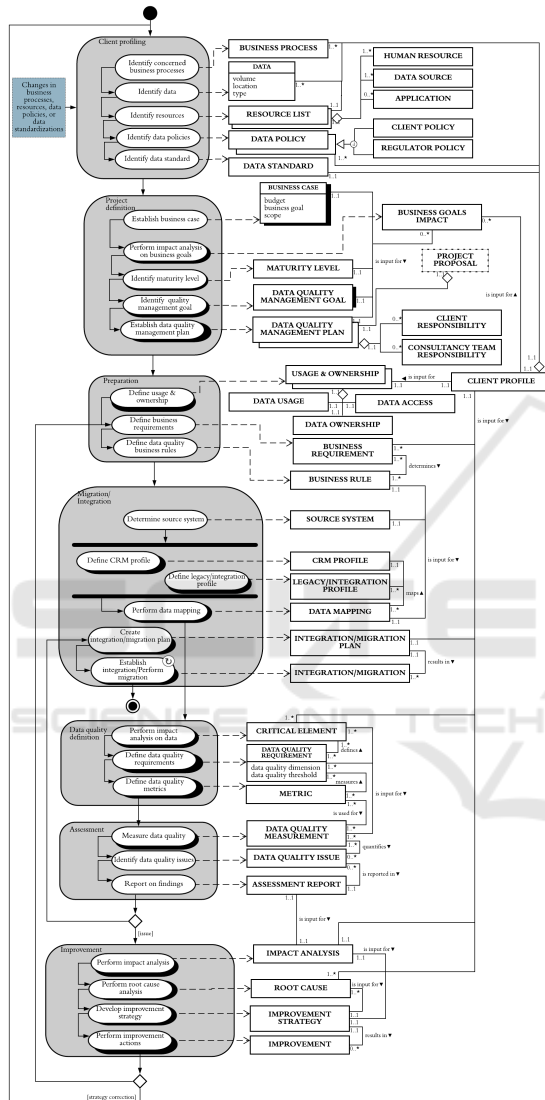


Figure 1: The CRM-DQMF.

Client Profiling is performed to reconstruct the client’s organisational environment with regards to the CRM solution. The output is a client profile which can be utilised for the definition of data roles, business impact analyses, and root cause detection. The *Project Definition* is executed once at the beginning of every project to indicate what the project will entail in terms of DQ management. Its output is a unique DQ management plan, which determines the utilisation of

the remainder of the CRM-DQMF. *Preparation* gathers the required information for the data mapping and assessment phases, which includes the definition of data roles, business requirements, and business rules. *Migration/Integration* is performed to migrate and/or integrate the data with the new CRM solution. *DQ Definition* and *Assessment* are performed to define and measure the DQ, and *Improvement* is performed to improve potential DQ issues. DQ is improved by refining business requirements and business rules till DQ is determined to be of sufficient quality. Then a migration/integration plan is established and executed. Once the migration or integration is established, the data lifecycle does not come to an end, hence DQ is still required to be managed. The phases *Project definition* and *Migration/Integration* are no longer part of the CRM-DQMF, and DQ monitoring will take place through continuous assessment and improvement of DQ. The framework is reconstructed as can be seen in figure 2. The distinct phases along with their activities are elaborated on below.

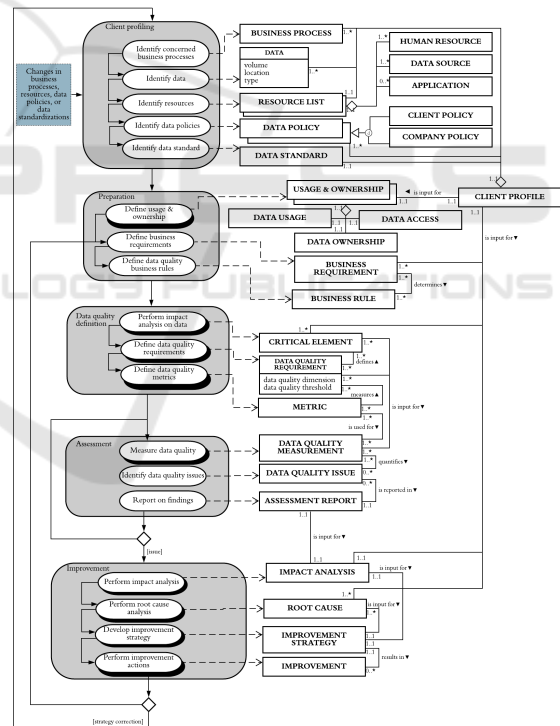


Figure 2: The CRM-DQMF after migration/integration is established.

3.2.1 Client Profiling

To understand and reconstruct the client’s organisational environment with regards to the CRM solution, a client profile is established containing information on the data, business processes, resources, data poli-

cies, and data standards. Below, the different activities that gather the required information are explained.

Identify Data: The concerned data is identified, so it is known which data should be subject to DQ management practices. The volume of the data is required in order to appropriately define a migration or integration strategy and to properly indicate the magnitude of potential DQ issues. The location of the data is required to properly indicate where the data affects the business, and hence determine the kind of DQ issues. The type of the data is required to determine the most appropriate DQ definition and measurement techniques.

Identify Concerned Business Processes: Business processes that are concerned with the CRM solution are identified. The business processes create, use, move, or modify the concerned data, and form a technical and business process landscape indicating the whereabouts and purposes of the data.

Identify Resources: The resources of the data are identified. This includes human resources, such as employees that enter the data, data sources that produce the data, and applications that utilise, move, or modify the data. The resources provide insights on the places of potential business impact caused by DQ issues. Subsequently, it can be used as input for the development of an improvement strategy.

Identify Data Policies: Data policies at the client's side are identified, as well as regulatory policies. They are directives that codify principles and management intent into rules that govern the data. Data policies might include, for example, rules about data classifications of criticality or GDPR. The data policies are input for the definition of DQ requirements.

Identify Data Standard: The existing data standard for all concerned data is identified. A data standard conditions the data to ensure that it meets rules for content and format. Data standards contribute to the definition of DQ, since they provide a means for comparison. The data standard requires continuous reviewing and refinement.

3.2.2 Project Definition

The project definition phase of the CRM-DQMF is the only phase that is executed by default for every project. As aforementioned, every client and project of a CRM D&C team is unique. Therefore, every

project requires its own DQ management plan. In the project definition phase of the CRM-DQMF, the DQ management plan is defined. This definition takes several actions, which are explained in the following paragraphs.

Establish Business Case: A CRM business case is established to define the CRM D&C project. This business case includes the business goals of the client and the CRM D&C project scope. The business goals and scope indicate whether data is required to be of high quality, and to what extent DQ management is of relevance. The budget of the client is determined, as this indicates the monetary boundaries of the CRM D&C project and the possible inclusion of the DQ management services of the CRM D&C team in the project proposal. Often, concessions have to be made either on DQ to deliver client experience within the constraints of costs and technique, or on the budget from the client's side.

Perform Impact Analysis on Business Goals: By defining the impact of poor DQ on the business goals, as well as the way that high DQ will enable the business goals the importance of DQ management is emphasized. This creates awareness on the topic for the client and makes an indication of the need for DQ management. When business impacts of poor DQ are defined to be negligible for the specific client and project, the remainder of the CRM-DQMF can be discarded. This might result in less effort by the CRM D&C team, as potential redundancy of the CRM-DQMF can be detected at an earlier point in time making the CRM-DQMF less expensive.

Identify Maturity Level: The DQ management maturity level of the client is taken into account for the design of a DQ management plan, as this indicates to what extent the client requires the assistance of a CRM D&C team in terms of DQ management. The activity is extracted from the insights on the DQ management expertise level of the client from the case study, and its relevance is supported by Spruit and Pietzka (Spruit and Pietzka, 2015).

Identify DQ Management Goal: Using the DQ management maturity matrix the client's goals of DQ management are indicated as well. As every client and project is unique, the goals of DQ management depend on the scope of the CRM D&C project and the business goals of the client, which determines the importance of DQ management. The current maturity level of DQ management next to the goal matu-

rity level indicates at which capabilities of DQ management the client requires to grow with regard to the CRM D&C project. This is used as input for the establishment of the DQ management plan.

Establish DQ Management Plan: Formatted by the business case, and using the input of the impact analysis, the current maturity level, and the goal maturity level of the client, the CRM D&C team develops a DQ management plan together with the client. The key activity is the definition of the key roles and responsibilities for the realisation of the target DQ management maturity level for each DQ management capability. It describes the DQ management services provided by the CRM D&C team, as well as the role descriptions and responsibilities of the client. The plan determines the remainder of the utilisation of the CRM-DQMF.

3.2.3 Preparation

The preparation phase gathers the required information for the data mapping and the assessment.

Define Usage & Ownership: For all data, the usage, ownership, and access are defined. Data usage includes all that make use of the data. Data ownership defines who is responsible for the data. Data access defines all that have access to the data.

Define Business Requirements: Business requirements are extracted from the business goals with the CRM D&C project. The business requirements describe what needs to be done to achieve the business goals.

Define DQ Business Rules: Business rules are defined and refined, describing expectations about the concerned data. They should be created through analysis of business processes, data policies, data standards, business impact of data, assessment reports, and common sense. Business rules are generally associated with the way data is collected or created. For example, when a client wants to send monthly newsletters to a specific sample of its customers as part of its marketing strategy in CRM, a business rule could be about the population of demographic fields such as *birth date*, or contact information fields such as *e-mail address*. In this case, a validity rule might describe the format of the field *birth date* in 'dd/mm/yy', and a completeness rule might describe the population of the field *e-mail address* to be mandatory.

3.2.4 Migration/Integration

To prepare for data migration or integration, the source system of the specific data is identified, as this determines the applicable data standards. Additionally, a mapping model is developed by defining the profile of the CRM platform and map this to the definition of the legacy CRM or the integration. After data mapping, the DQ is assessed. Once the DQ is assessed and, when required, improved, a migration and/or integration plan is created and executed. This is done after (some iterations of) assessment and improvement practices, since ideally potential issues regarding DQ are resolved before migration or integration is established to prevent more significant problems. Most issues concerning DQ are discovered when the migration is performed in a test environment as part of a dry run. The rotating arrow is added as an extension on the PDD notation as introduced by van de Weerd and Brinkkemper (van de Weerd, 2009). It indicates that migration practices can be performed multiple times as dry runs. When the issues as discovered by a dry run are eliminated, the migration will be established either again as a dry run, or in production. After the migration has been established, the DQ definition and assessment are required to be ongoing processes, hence the process loops back to DQ definition in case of no DQ issues till the end of the data life cycle. When DQ issues occur, improvement practices are implemented first.

3.2.5 DQ Definition & Assessment

The input for the DQ assessment phase is the gathered knowledge of the previous phases. The output is a DQ report.

Perform Impact Analysis: By performing an impact analysis, critical data elements are identified. Critical data elements represent data that is of utmost importance for the achievement of the business goals. Those elements are required to comply with their DQ definitions. The result of the impact analysis is a prioritised list of data elements which can be used by the team to focus their work efforts.

Define DQ Requirements: DQ is defined by means of DQ dimensions and DQ thresholds. DQ dimensions enable the characterization of rules (e.g., e-mail address must be populated) and findings (e.g., e-mail address is 98% complete). They facilitate a mutual understanding of what is being measured. The DQ dimensions provide the basis for the definition of meaningful metrics. The DQ threshold defines the requirement belonging to the DQ dimension.

Define DQ Metrics: Once the DQ dimensions are defined, metrics can be defined in order to quantify the findings of DQ. For example, a DQ business rule can be for the field *e-mail address* to be mandatory, which translates into the DQ dimension *completeness*. The metric that can be used to measure the completeness of the field *e-mail address* can be of type ratio, dividing the number of records where the field is populated by the total amount of records, and multiply this by 100 to get the percentage of complete records.

Measure DQ: DQ is measured either subjectively or objectively (Cichy and Rass, 2019). The metrics are used to quantify the measurements. The output is the quantified measurements of DQ.

Identify DQ Issues: Based on the measurements and DQ business rules, DQ issues are identified. DQ issues are identified by setting status indicators for all data in terms of its dimension(s) and thresholds. For example, the status indicator of the dimension of completeness for the field *e-mail address* can be indicated *Unacceptable* when the measurement results in the threshold of below 80% complete.

Report on Findings: The final output is an assessment report of the DQ and potential issues. The assessment report might offer a new perspective on the concerned data, from which new business rules could be articulated. When DQ issues occur, improvement practices will take place.

3.2.6 Improvement

The improvement phase of the CRM-DQMF is only executed when DQ issues are reported on in the output of the assessment phase. When improvement activities have been applied, the CRM-DQMF loops back to the *Preparation* phase to review business requirements and business rules. In case of a strategy correction as part of the improvement strategy, the *Client profiling* phase should be revisited to review the organisational environment.

Perform Impact Analysis: The identified DQ issues are quantified and prioritized based on business impact. Business impacts include monetary costs of poor DQ, as well as non-monetary impacts. It also takes into account the criticality of the data, the volume of the data, the number of business processes and stakeholders impacted by the issue, and the risks associated with the issue. This information is all extracted during the *Client profiling* and *Preparation* phases of

the framework. The output is a ranked list of DQ issues that should be taken into account within the improvement strategy.

Perform Root Cause Analysis: Ideally, the DQ issues are remediated at their root cause (Batini et al., 2009). This could also mean controls and process improvements to prevent further DQ issues from happening. Therefore, a root cause analysis is performed to identify the root causes of DQ issues.

Develop Improvement Strategy: Based on the impact analysis, an improvement strategy is developed, evaluating the costs of the issue against the costs of the required improvement actions. The improvement strategy ranks the issues that can be addressed immediately and at low costs, as well as more strategic improvements, such as root cause remediation and prevention practices. It contains improvement goals that are specific, achievable, and based on a quantification of the business impacts.

Perform Improvement Actions: The improvement strategy is put into practice. This might result in revisiting *Client profiling* or *Preparation* practices, or direct improvements in the data. Either way, assessment is performed again to assess the DQ.

4 VALIDATION

To validate the CRM-DQMF, a validation model of the artifact is drafted. The validation model, or design theory, consists of a description of the properties of the artifact and the interaction with the problem context (Wieringa, 2014). The discussions that form the design theory are facilitated by the use of confirmatory focus groups (Tremblay et al., 2010) with a total of 6 experts (see table 3). Additionally, to observe and measure how well the CRM-DQMF supports DQ management in CRM D&C projects, the evaluation model of Moody (Moody, 2003) and evaluation criteria by Prat et al. (Prat et al., 2015) are utilised. The constructs of desired qualities (*Perceived Alignment with CRM D&C* and *Perceived Effectiveness*), *Perceived Ease of Use*, *Perceived Usefulness*, and *Perceived Completeness* are measured by means of an interactive questionnaire, of which the results on a 5-point Likert scale can be found in table 4.

The main insights that are extracted are elaborated on in the following paragraphs.

Table 3: Participants Validation.

ID	Role	Years of Experience
P1	Consultant	4
P2	Senior Analyst	2
P3	Manager	9
P4	Senior Consultant	4
P5	Senior Consultant	7
P6	Senior Analyst	2

Table 4: Participants Validation.

Construct	P1	P2	P3	P4	P5	P6	Average
Alignment with CRM D&C Business	4,5	5	4,5	4,5	4	3,5	4,3
Perceived Effectiveness	5	5	4,5	4,5	3,5	5	4,6
Perceived Completeness	4,5	3	4	2,5	3	3	3,8
Perceived Ease of Use	4,5	3,5	4	4	2,5	3	3,6
Perceived Usefulness	5	4,5	4,5	4,5	3,5	5	4,5
Intention to Use	5	5	4	4,5	3	4,5	4,3

Expensiveness: The establishment of a DQ management plan as it is presented in the first design of the CRM-DQMF might be too expensive for smaller projects where small amounts of data are involved. To mitigate this problem, the establishment of the DQ management plan should be more integrated in the creation of the general CRM D&C project proposal, rather than an activity on its own. However, this is not part of the CRM culture at the case study environment yet, and thus this might require change management practices.

Awareness: Clients might not be aware of the importance of DQ management, hence do not want to spend their budget on DQ management services of the CRM D&C team. Therefore, the impact of poor quality data on the business goals needs to be defined at the start of a project, as well as the way that high quality data will enable the business goals (International DAMA, 2017). This emphasizes the importance of DQ management, creating awareness of the topic and making an indication of the need for DQ management. However, this needs to be done as inexpensive as possible in order to be profitable for the CRM D&C team, as this influences the development

of the DQ management plan and thus the project proposal.

Agility: The agile project approach of CRM D&C projects did not come through cogent enough according to the validation sessions. The participants argued that they would like the CRM-DQMF to guide the CRM D&C team in integrating DQ management in the concept of so-called *sprints* in agile projects, where specific work is selected for a set period of time. First thoughts on this matter indicate that the CRM-DQMF is supposed to support an ongoing process, which could be translated into sprints, where the assessment and potential improvement of DQ iteratively takes place in every new sprint in the project.

High-level: The contemporary design of the CRM-DQMF is deemed too high level to put into practice as it is. The CRM-DQMF creates awareness on the importance of DQ management amongst CRM D&C teams and provides relevant insights in what DQ management encompasses, rather than providing a step-by-step guidance in implementing DQ management in CRM D&C projects.

5 DISCUSSION

The factors that form a potential threat to the validity of this research are elaborated on in following sections, along with the possible diminishing of those threats. Subsequently, the limitations to this research are discussed.

5.1 Threats to Validity

The five types of validity as described by Johnson (Johnson, 1997) are used to examine the validity of this research.

Descriptive Validity: For this research, only one researcher interviewed participants and conducted documentation analysis, which eliminates the possibility to achieve this validity type through investigator triangulation. To improve the descriptive validity of this research nonetheless, all conducted interviews and validation sessions were recorded to facilitate more accurate recalls of the researcher.

Interpretive Validity: To ensure this type of validity, the researcher regularly incorporated participant feedback within the case study (Johnson, 1997). This is done through the utilisation of the question type

interpreting questions as proposed by Kvale (Kvale, 1996) to inspect whether the interviewee's answer is interpreted correctly.

Theoretical Validity: To achieve theoretical validity, fieldwork is incorporated in this research. This fieldwork consists of several elements: participation in training session which were facilitated to CRM D&C consultants; close observation of a collaboration tool used by the CRM D&C community at the case company; attending presentations on the execution of specific CRM D&C projects; joining a day of scrum meetings between the case company and a client of the financial industry; and gain certificates which are required/recommended for CRM D&C employees.

Internal Validity: Method triangulation is utilised to achieve internal validity (Kaplan and Maxwell, 2005; Johnson, 1997). This means that more than one research method is used, namely literature review, expert interviews, and documentation analysis. Subsequently, data triangulation is applied by making use of multiple data sources (Johnson, 1997). Multiple expert interviews are conducted with participants from varying backgrounds. Subsequently, documentation from a variety of sources is examined for the documentation analysis.

External Validity: The case study is performed at one organisation. However, various literature sources are used for the design of the CRM-DQMF, and documentation utilised for the documentation analysis originated from two additional organisations. Furthermore, the abstractness and high-level approach of the CRM-DQMF increases its external validity, as it facilitates generalisability to CRM D&C projects varying in client, industry, and business goals.

5.2 Limitations

First of all, due to time and resource restrictions, this research was not able to investigate the actual adoption of the CRM-DQMF within a CRM D&C project. Consequently, all conclusions are extracted from non-empirical sources, and the experience and data of experts.

Second, most participants were not consciously familiar with DQ management practices. Therefore, much effort went to exploring the expertise level of participants and being able to conduct the interviews in such a way that participants understood the concepts, while phrasing interview questions to extract

required information without creating researcher bias. Additionally, the expert interviews with the participants evolved over time. This contributes to the formulation of some question types, such as follow-up and probing questions (Kvale, 1996), and the answers to those questions weigh more when they are agreed with by other interviewees. The disadvantage is that it could create bias, as interviewees might have been pushed towards a certain direction. To mitigate this disadvantage, results from previous interviews were only provided when the interviewee already provided an answer on their own, and there existed sufficient grounds for suspecting the previous results might be applicable for the current interview as well.

Third, potential participants for the validation sessions had busy and asynchronous agendas, which made it difficult to schedule focus groups of sufficient sizes. In the end, there was chosen to perform mini focus groups consisting of two participants (Nyumba et al., 2018).

Last, the research process has been impacted by the need for all research efforts to be arranged online due to the COVID-19 regulations set by the government, the university, and the case company. This might have resulted in less sufficient sampling results. Subsequently, it might have influenced the interpretation of the researcher, as online settings were sometimes lacking in terms of connection.

6 CONCLUSIONS AND FUTURE RESEARCH

6.1 Conclusions

The designed CRM-DQMF combines scientific literature and practitioner's insights on DQ management and CRM D&C. It provides a high-level overview of DQ management practices incorporated in CRM D&C projects. With its current design, the CRM-DQMF is a tool to plan on opportunities for the incorporation of DQ management in CRM D&C projects. This incorporation involves the recognition of variety in clients and projects by the establishment of a unique DQ management plan. This plan describes to what extent DQ management services of the CRM D&C team are required for the specific project. The CRM-DQMF contains the following components:

- *Client profiling* to gather required knowledge for DQ management and the CRM solution;
- *Project definition* to establish the data quality management plan for the project;
- *Preparation* to define data roles and business requirements;

- *Migration/Integration* to establish migration and/or integration;
- *DQ definition* to define quality and metrics;
- *Assessment* to measure DQ and report on findings; and
- *Improvement* to define an improvement strategy and improve DQ.

6.2 Future Research

This research leaves multiple opportunities for further research. First, improvement opportunities can be found in *perceived ease of use* and *perceived completeness* based on the validation. Thus, further research could focus on the investigation of the usability and exhaustiveness of the CRM-DQMF. The contemporary CRM-DQMF is a high-level overview of DQ management practices incorporated in CRM D&C projects. However, as found within this study, to be of optimal use for CRM D&C teams, it requires more step-by-step guidance on how to perform or consult on the individual activities. Hence, more research can be done on the incorporation of step-by-step guidance on the execution of each activity that is included in the CRM-DQMF as well as the utilisation of the CRM-DQMF as a whole, to ensure its usability for CRM D&C teams. This might include further research to investigate on how CRM D&C teams can be directed on which components of the CRM-DQMF suit their client's situation best.

Second, further empirical validation of this research is required. This research solely evaluates the design through validation sessions using a design theory and a questionnaire. Possible evaluation can include, for example, expert interviews, technical action research at actual CRM D&C projects, or surveys. Furthermore, the sampling results of the case study can be extended to experts from other fields to include additional perspectives next to those that are utilised within this study. The results can then be used for the improvement of the CRM-DQMF.

Third, this research assumes that the participants of the case study are able to decide which practices would best fit the needs of the CRM D&C projects of the organisation. During the case study and validation sessions, some experts argued that they would require more guidance or persuasion for the application of the CRM-DQMF or DQ management practices at all. DQ management is not part of the CRM D&C culture within this case study. Hence, the adoption of the CRM-DQMF might involve change management to ensure CRM D&C teams involve DQ management practices within CRM projects more explicitly, which could result in better CRM solutions. Fourth, the implementation of the CRM-DQMF might be too ex-

pensive as it is. Therefore, an opportunity for further study lies in how to integrate the establishment of a DQ management plan into the development of a CRM D&C project proposal in an as inexpensive as possible manner, as this is only implicitly mentioned within this research.

Lastly, potential further research can review the incorporation of the CRM-DQMF into agile project approach practices, such as sprints, as this is only implicitly mentioned within this research.

REFERENCES

- Ali, H., Moawad, R., and Hosni, A. (2017). A cloud interoperability broker (CIB) for data migration in SaaS. *Future Computing and Informatics Journal*. 10.1016/j.fcij.2017.03.001.
- Batini, C., Barone, D., Mastrella, M., Maurino, A., and Ruffini, C. (2007). A framework and a methodology for data quality assessment and monitoring. *Proceedings of the 12th International Conference on Information Quality*.
- Batini, C., Cappiello, C., Francalanci, C., and Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM Comput. Surv.*, 41(3). 10.1145/1541880.1541883.
- Budgen, D. and Brereton, P. (2006). Performing systematic literature reviews in software engineering. *ICSE '06 Proceedings of the 28th International Conference on Software Engineering*. 10.1145/1134285.1134500.
- Buttle, F. and Maklan, S. (2019). *Customer Relationship Management: Concepts and Technologies*. Routledge, 4th edition.
- Cichy, C. and Rass, S. (2019). An overview of data quality frameworks. 7:2169–3536. 10.1109/ACCESS.2019.2899751.
- Cricelli, L., Famulari, F., Greco, M., and Grimaldi, M. (2019). Searching for the one: Customer relationship management software selection. *Journal of Multi-Criteria Decision Analysis*, pages 173–188. 10.1002/mcda.1687.
- Cruz-Jesus, F., Pinheiro, A., and Tiago, O. (2019). Understanding CRM adoption stages: empirical analysis building on the TOE framework. *Computers in Industry*, 109:1–13.
- Davenport, T. and Harris, J. (2017). *Competing on Analytics: Updated, with a New Introduction: The New Science of Winning*. Harvard Business Review Press.
- Gill, A. and Henderson-Sellers, B. (2006). Measuring agility and adaptability of agile methods: a 4 dimensional analytical tool. *The IADIS international conference on applied computing*.
- Hanson, G. (2020). State of CRM data management 2020 report. <https://www.salesforceben.com/the-state-of-crm-data-in-2020-our-highlights-from-300-organizations/>.

- Heinrich, B., Hristova, D., Klier, M., Schiller, A., and Szubartowicz, M. (2018). Requirements for data quality metrics. *J. Data and Information Quality*, 9(2):45–61. 10.1145/3148238.
- IDC (2020). Global datasphere. IDC. https://www.idc.com/getdoc.jsp?containerId=IDC_P38353#:~:
- International DAMA (2017). *DAMA-DBoK: Data Management Body of Knowledge*, volume 2. Technics Publications. 9781634622363.
- Johnson, R. (1997). Examining the validity structure of qualitative research. *Education*, 141.
- Kaplan, B. and Maxwell, J. (2005). *Qualitative Research Methods for Evaluating Computer Information Systems*. Springer. 10.1007/0-387-30329-4.2.
- Kvale, S. (1996). Interviews: An introduction to qualitative research interviewing. *Thousand Oaks*.
- Laney, D. B. (2018). *Infonomics: How to monetize, manage, and measure information as an asset for competitive advantage*. Taylor & Francis Group.
- Levy, Y. and Ellis, T. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Information Science Journal*, 9:181–212. 10.1049/cp.2009.0961.
- Missi, F., Alshawi, S., and Fitzgerald, G. (2005). Why crm efforts fail? a study of the impact of data quality and data integration. *Hawaii International Conference on System Sciences*, 38. 10.1109/HICSS.2005.695.
- Moody, D. (2003). The method evaluation model: a theoretical model for validating information systems design methods. *ECIS 2003 proceedings*, 79.
- Nagle, T., Redman, T., and Sammon, D. (2020). Assessing data quality: A managerial call to action. *Business Horizons*, 63:325–337. <https://note.org/10.1016/j.bushor.2020.01.006>.
- Nyumba, O., Wilson, K., Derrick, C., and Mukherjee, N. (2018). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution*, 9. 10.1111/2041-210X.12860.
- Otto, B., Wende, K., Schmidt, A., and Osl, P. (2007). Towards a framework for corporate data quality management. *Australasian Conference on Information Systems*, pages 916–926.
- Pipino, L., Lee, Y., and Wang, R. (2002). Data quality assessment. *Communications of the ACM*, 45(4). 10.1145/505248.506010.
- Prat, N., Comyn-Wattiau, I., and Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, 79(3):229–267. 10.1080/07421222.2015.1099390.
- Robson, C. and McCartan, K. (2016). *Real world research*. John Wiley & Sons, 4th edition.
- Rosing, M. and Gill, A. (2015). Applying agile principles to bpm. In *The complete business process handbook*, volume 1, pages 553–577. Elsevier - Morgan Kaufman. 10.1016/B978-0-12-799959-3.00027-6.
- Spruit, M. and Linden, V. v. d. (2019). BIDQI: The business impacts of data quality interdependencies model. *Department of Information and Applied Sciences, Utrecht University*.
- Spruit, M. and Pietzka, K. (2015). MD3M: The master data management maturity model. *Computers in Human Behavior*, 51:1068–1076. <http://dx.note.org/10.1016/j.chb.2014.09.030>.
- Stone, M., Aravopoulou, E., Gerardi, G., Todeva, E., and Weinzierl, L. (2017). How platforms are transforming customer information management. *The Bottom Line*, 30(3):216–235. 10.1108/BL-08-2017-0024.
- Thalheim, B. and Wang, Q. (2013). Data migration: A theoretical perspective. *Data & Knowledge Engineering*, 87:260–278.
- Tremblay, M., Hevner, A., and Berndt, D. (2010). The use of focus groups in design science research. pages 1221–143. 10.1007/978-1-4419-5653-8_10.
- van de Weerd, I. (2009). Advancing in software product management: An incremental method engineering approach. <http://igitur-archive.library.uu.nl/dissertations/2009-0826-200144/UUindex.html>.
- van de Weerd, I. and Brinkkemper, S. (2008). Meta-modeling for situational analysis and design methods. chapter III, pages 35–54. 10.4018/978-1-59904-887-1.ch003.
- Wang, R. (1998). A product perspective on total data quality management. *Communications of the ACM*, 41(2):58–65.
- Wieringa, R. (2014). *Design Science Methodology: for Information Systems and Software Engineering*. Heidelberg: Springer. 10.1007/978-3-662-43839-8.
- Yin, R. K. (2003). *Case Study Research: Design and Methods (3rd)*. SAGE Publications.
- Zahay, D., Krishen, A., and Peltier, J. (2012). Building the foundation for customer data quality in CRM system for financial services firms. *Journal of Database Marketing & Customer Strategy Management*, 19(1):5–16. 10.1057/dbm.2012.6.

APPENDICES

- Data Quality Management Methods, see: <https://osf.io/xkquj/>
- Interview Protocol, see: <https://osf.io/3c9ys/>
- CRM-DQMF Design, see: <https://osf.io/375up/>