Customer Data Management

MAHDIS SEHAT
RENÉ PAVEZ FLORES

Master of Science Thesis
Stockholm, Sweden 2012
Customer Data Management

Mahdis Sehat
René Pavez Flores

Master of Science Thesis INDEK 2012:89
KTH Industrial Engineering and Management
Industrial Management
SE-100 44 STOCKHOLM
Abstract
As the business complexity, number of customers continues to grow and customers evolve into multinational organisations that operate across borders, many companies are faced with great challenges in the way they manage their customer data. In today’s business, a single customer may have a relationship with several entities of an organisation, which means that the customer data is collected through different channels. One customer may be described in different ways by each entity, which makes it difficult to obtain a unified view of the customer. In companies where there are several sources of data and the data is distributed to several systems, data environments become heterogenic. In this state, customer data is often incomplete, inaccurate and inconsistent throughout the company. This thesis aims to study how organisations with heterogeneous customer data sources implement the Master Data Management (MDM) concept to achieve and maintain high customer data quality. The purpose is to provide recommendations for how to achieve successful customer data management using MDM based on existing literature related to the topic and an interview-based empirical study. Successful customer data management is more of an organisational issue than a technological one and requires a top-down approach in order to develop a common strategy for an organisation’s customer data management. Proper central assessment and maintenance processes that can be adjusted according to the entities’ needs must be in place. Responsibilities for the maintenance of customer data should be delegated to several levels of an organisation in order to better manage customer data.

Keywords: Customer Data Management, Master Data Management, Customer Data Quality, Data Quality Management.
Preface

This report is our Master Thesis for the conclusion of our Master program at the institution of Industrial Management and Engineering at the Royal Institution of Technology. We would like to thank the people of Scania’s Franchise Standards & Volume Planning department for a pleasant time at their offices and for their help. We would also like to thank all the interviewees at Scania, Ernst & Young, Xlent and DeLaval for their expertise and enthusiasm. We would especially like to thank our supervisor Per-Erik Anderson and our assignor Daniel Boëthius for the opportunity to write our thesis in collaboration with Scania and for a great experience. We would also like to thank our supervisor Jannis Angelis and examiner Mats Engwall for their guidance.

Stockholm, August 2012

Mahdis Sehat René Pavez Flores
# Table of Content

1. **Introduction** .......................................................................................................................... 6
   1.1 Background ............................................................................................................................ 6
   1.2 Aim ....................................................................................................................................... 7
   1.3 Research Questions .............................................................................................................. 7
   1.4 Delimitations ........................................................................................................................ 8
2. **Methodology** ........................................................................................................................... 9
   2.1 Research Approach .............................................................................................................. 9
   2.2 Research Process .................................................................................................................. 10
   2.3 Literature Study ................................................................................................................. 11
   2.4 Empirical Study ................................................................................................................... 12
   2.5 Validity, Reliability and Generalisability ........................................................................... 16
3. **Literature Study** .................................................................................................................... 18
   3.1 Customer Data Management .............................................................................................. 18
   3.2 Crucial Factors and Challenges .......................................................................................... 19
   3.3 Change Management ......................................................................................................... 20
   3.4 Fundamentals of Customer Data Management .................................................................. 22
4. **Empirical Study** .................................................................................................................... 33
   4.1 Empirical Setting .................................................................................................................. 33
   4.2 Current Customer Data Management at Scania ................................................................. 36
   4.3 Current Customer Data Management at DeLaval ............................................................... 38
   4.4 Customer Data Management ............................................................................................. 39
   4.5 Crucial Factors and Challenges .......................................................................................... 43
   4.6 Fundamentals of Customer Data management ................................................................... 47
5. **Discussion** ............................................................................................................................. 55
   5.1 Customer Data Management .............................................................................................. 55
   5.2 Crucial Factors and Challenges .......................................................................................... 56
   5.3 Data Governance ................................................................................................................ 58
   5.4 Data Stewardship .............................................................................................................. 60
   5.5 Data Quality Management ................................................................................................. 61
   5.6 Data Quality Assessment and Improvement ...................................................................... 64
6. **Conclusion** ............................................................................................................................. 66
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Conclusion</td>
<td>66</td>
</tr>
<tr>
<td>6.2 Recommendations</td>
<td>67</td>
</tr>
<tr>
<td>6.3 Limitations and Future Research</td>
<td>69</td>
</tr>
<tr>
<td>References</td>
<td>71</td>
</tr>
<tr>
<td>Appendix</td>
<td>74</td>
</tr>
<tr>
<td>Appendix A: Wordlist</td>
<td>74</td>
</tr>
</tbody>
</table>
1. Introduction

This master thesis is about how the concept of Customer Data Management can be implemented in an organisation with heterogeneous customer data sources and aims to generate recommendations for how successful customer data management is achieved. The introduction begins with a background, giving the reader a basic knowledge about the topic being researched and why it is of interest. Furthermore, the aim and research questions of the paper are presented.

1.1 Background

As the business complexity, number of customers, number of lines of business, and number of sales and service channels continue to grow, many organisations have evolved into a state with many customer data sources and systems for managing the data. This raises great concern and presents challenges regarding customer data management. (Berson, Dubov 2007)

In companies where there are several sources of data and the data is distributed to several systems, data environments become heterogenic with different systems, data models and processes being used to manage data within the company. (Batini, Scannapieca 2006) In this state, customer data is often incomplete, inaccurate and inconsistent throughout the company (Berson, Dubov 2007). The key issues for managing data are poor data quality and unclear definitions of the data collected. Other issues involve inadequate processes for maintaining data, unclear data ownership and absence of continuous data quality maintenance. (Silvola et al. 2011)

Every company deals with customers and within the company, each customer may have a relationship with several entities: marketing, sales, support, maintenance, customer satisfaction, billing or service. The customer may be described with different aspects of the customer’s attributes in each unit. Furthermore, every business application may value the attributes differently and define data quality in different ways depending on the business context. For example, the telemarketer wants to avoid calling the same customer twice and therefore values accuracy of telephone number highly, while shipping is more concerned with location information. (Loshin 2009) The requirements for successful data management are; a well defined data model, clear ownership and responsibilities definitions, constant data monitoring and maintenance, organisational structure that supports the processes involved, managerial support and information systems that utilise the unified data model. (Silvola et al. 2011)

Customer Data Management (CDM) is a term evolved from the concept Master Data Management (MDM), where focus lies on managing customer data (Berson, Dubov 2007). MDM is a major research area that aims to allow users to access data through a unified view, though the data is stored in heterogeneous data sources (Batini, Scannapieca 2006). The essence of MDM is to organise an enterprise’s view of the key business information objects
and govern their use and quality to achieve operational efficiency and open opportunities for optimisation. The intention of an MDM program is to create a single storage of high quality master data that then feeds data across the organisation with a synchronised and consistent view of enterprise data. (Loshin 2009) Master data is data that is used across multiple business units (Berson, Dubov 2007). The MDM concept, and thereby the CDM concept, concentrates on the collection and maintaining of high quality data through standardised maintenance processes and clear data ownership (Silvola et al. 2011).

Having accurate and centralised customer data benefits the sales department and increase revenue by allowing the organisation to better gain insight into their customers’ objectives, demands and tendency to request additional products and services. In addition, customer satisfaction can be improved and loyalty increased by achieving a complete picture of the customer, which enables the firm to offer customised products and services. Having centralised customer data is simpler to maintain due to the fact that there is only one single version of the data, which also reduces costs. (Berson, Dubov 2007)

1.2 Aim

The aim of this thesis is to study customer data management in an organisation with heterogeneous customer data sources and to generate recommendations for methods to achieve successful customer data management based on a case study, Scania, existing literature related to the topic, expertise knowledge gathered from interviews with consultants and a benchmarking company, DeLaval.

The issue is inconsistent customer data caused by heterogeneous customer data management and the nature of today’s multinational customers as well as undefined standards and responsibilities in the maintenance and data collection process for the organisation.

1.3 Research Questions

In order to reach the aim of this thesis, the authors seek to find answers to the following research questions:

- How is successful customer data management achieved in an organisation using the Master Data Management concept?
- What are the critical factors and challenges of customer data management?
- How is high data quality achieved and maintained in a Master Data Management environment?
1.4 Delimitations

The thesis is delimited to customer data management according to the concept of Master data Management. This thesis is delimited to not cover the technological aspect as the essential factors and challenges regarding Master Data Management are concerned with organisational aspects. Therefore, the literature study does not cover implementing of technologies for customer data management nor the challenges of it. The fundamentals of customer data management are delimited to the terms Data Governance, Data Stewardship, Data Quality Management, and Data Quality Assessment and Maintenance. The choice of relevant terms for the thesis is based on the crucial factors and challenges found in the literature study. Further, theories concerning Change Management are included since the issues revealed by the literature stress organisational challenges regarding achieving acceptance of changes.

According to the literature, Master Data Management is a concept that aims to provide a unified view of the data, though the data sources are heterogeneous. The case study company, Scania, is chosen based on their request to achieve a unified view of their customers despite the fact that customer data is collected from many sources. Due to the time limit of the thesis, interviews were limited to central entities and a dealer on the Swedish market at the case company, Scania. Since the topic of customers and customer data is a delicate issue, the writers of this thesis have experienced difficulties of finding companies willing to participate in interviews and share information regarding their customer data management, which has resulted in one benchmarking company. Due to the delimitations mentioned, the empirical study is used to achieve a general understanding of the management of customer data at the companies. The thesis does not cover processes and activities that should be used to achieve high data quality, since this should be based on the business processes, which requires a better understanding of the company and its processes.
2. Methodology

In the methodology section, the research approach and research process chosen for the thesis is discussed. Additionally, the method used for the selection of literature and empirical objects, and the validity and generalisability of the thesis are described. The thesis is based on two studies: literature study and empirical study. The literature study covers literature regarding Customer Data Management, Master Data Management, Data Governance, Data Stewardship, Data Quality Management and Data Quality Assessment and Improvement. The empirical study is based on a case company, Scania; external expertise in form of opinions of consultants and one benchmarking company, DeLaval.

2.1 Research Approach

A research paradigm is a philosophical framework that guides how scientific research should be conducted (Collis, Hussey 2003). Collis and Hussey (2003) discuss two main paradigms, namely positivism and interpretivism. Positivism is a research paradigm that originated from the natural sciences. Its foundation is the assumption that social reality is singular and objective, and is not affected by the act of investigating it and therefore the researcher is seen as independent of what is being researched. The research involves a deductive process with a view to providing explanatory theories to understand social phenomena. (Collis, Hussey 2003)

Positivism and Interpretivism

In the positivistic paradigm, theories are used as the basis of explanation the occurrence of and to anticipate phenomena and predict their occurrence. Quantitative methods of analysis are characteristic for positivistic studies, as it is believed that social reality is measurable (Collis, Hussey 2003). Since the researcher is detached from the researched phenomena, the axiological assumption is that the research is value free and un-biased. The language used in the research is formal and in written in the third person (Cresswell 1998).

In contrast to the positivistic assumption on social reality, the interpretivism approach rests on the assumption that social reality is in our minds, it is subjective and multiple. Therefore, the social reality is affected by the researchers interactions with the phenomenon that is being investigated (Cresswell 1998). Research in this paradigm involves an inductive process with a view to providing interpretive understanding of social phenomena within a particular context. Reality is subjective and researchers must acknowledge that the research affected by their own values is biased. The language of the research is informal and uses a personal voice. (Collis, Hussey 2003) Strauss and Corbin (1990) came to the broad conclusion that interpretive research is any type of research where the findings are not derived from the statistical analysis of quantitative data.

Deduction and Induction

Researchers aim to formulate theories that give as accurate knowledge about reality as possible. Empirics are often used as the base for theories. The work for researchers lies in the ability of how to relate theory and reality with each other (Patel, Davidson 2011).
Research is deductive under a positivist paradigm (Collis, Hussey 2003). Deduction is defined as: “In logic, a rigorous proof, or derivation, of one statement (the conclusion) from one or more statements (the premises) – i.e., a chain of statements, each of which is either a premise or a consequence of a statement occurring earlier in proof.” (Britannica, A)

Deductive research describes a study in which a conceptual and theoretical structure is developed which is then tested by empirical observation (Collis, Hussey 2003). The reason for the use of deductive research is to attain better understanding for the matter, and create a structure for the study in order to eliminate redundant data collection during the empirical study.

Induction is defined as: “In logic, method of reasoning from a part to a whole, from particulars to generals, or from the individual to the universal” (Britannica, B). Inductive research is when theory is developed from observations of empirical reality. The inductive method is the opposite of the deductive method. The researcher creates a general from individual instances. (Collis, Hussey 2003)

Positioning

The writers of the thesis make the axiological assumption that they are detached and independent of the studied phenomenon, which is characteristic for positivism. However, the methodology of the research involves a number of interviews in order to obtain different perceptions of the problem and produce rich subjective qualitative data, a method, which is much related to an interpretivistic research approach. The analysis is then used to understand the situation. The methods used in this thesis are mostly associated with interpretivism. Collis and Hussey (2009) suggest that positivism and interpretivism should be regarded as two extremities of a continuum of paradigms where different paradigms can exist simultaneously. The conclusion of this thesis has been reached through both deduction and induction, as the authors first used deduction to create a theoretical foundation in order to examining the case studies after which a conclusion was reached through induction and overall the thesis is closer to interpretivism than positivism.

2.2 Research Process

Initially, research questions and aim for the thesis have been defined through collaboration with the thesis assignor and using existing literature. An initial literature study was done to understand the concept of MDM and Data Quality Management. The essential factors found in the initial literature study have been the base for the development of the interviews conducted for the empirical study. The literature study was then followed by an empirical study based on interviews at the case company, two consulting firms and one benchmarking company. As new topics were discovered during the empirical study, the literature study was then reviewed along the research process and additional literature study was done with focus on the essential factors and issues found through the initial literature study and empirical study. Additional telephone interviews have been conducted to supplement the initial interviews. The findings from both studies have been compared in an analysis after which a conclusion has been reached. The research process is illustrated in figure 2.1
2.3 Literature Study

Literature refers to the existing knowledge involving the researched phenomenon, and the literature research is a systematic process of identifying this existing knowledge. It is desired that as many relevant publications as possible are collected and read. The literature research increases the researchers’ knowledge about the researched topic and enables a foundation for the empirical study. (Collis, Hussey 2003) Literature refers to all collected secondary data. The first step is to define the scope of the literature research regarding time, geography and disciplinary approach. Further, it is important to define what sources of information is relevant for the study. The next step is to identify key words associated with the research topic. Once the literature is collected, it is time to analyse the data. Here, a thematic analysis is used where the themes of the relevant literature are categorised, and broken down into sub-groups. This enables structuring of the literature study. (Collis, Hussey 2003)

The literature is collected from libraries, databases and the Internet, and covers the concepts of Master Data Management, Customer Data Management, Data Governance, Data Stewardship, Data Quality Management and Data Quality Assessment and Improvement.

Master Data Management is the main topic of the thesis. Loshin was chosen as the primary source for MDM as the articles written by Loshin was considered highly relevant for the research questions and the author is considered to an industry thought-leader and a very recognised expert in information management (Search Data Management). Further, Cervo and Allen’s book (2011) was chosen as one of the main books in the literature study. This book emphasises on the practical aspects of implementing and maintaining MDM, and focuses on customer MDM (Cervo, Allen 2011), which is highly relevant for this thesis.

After extensive Internet research based on keywords related to the issue at hand, Batini’s, together with Scannapieca’s work appeared as highly relevant for this study in the category of Data Quality Management and Data Quality Assessment and Improvement. Articles co-written by Batini were chosen because besides being relevant, the authors of the publications made an extensive research on several different data quality assessment and

---

Figure 2.1: Illustration of the Research Process
present a good overview of these. Other publications were used to complement the theories of above mentioned authors in order to have an as non-biased theory as possible.

2.4 Empirical Study

The empirical study is based on three parts: interviews with stakeholders at a case company, Scania; interviews with consultants with expertise in Information Management and MDM; and an interview at a company that has implemented a central system for customer data management.

Scania has been chosen as the case company for this thesis since their issues and desires match well with the theories of MDM. Therefore, interviews have been conducted at Scania to understand the current customer data management and what the stakeholders want to achieve in the future. Interviews with consultants from Ernst & Young and Xlent were conducted at the firms’ Stockholm offices. The purpose of the interviews has been to get experts’ views on the matter. Both firms are consulting firms that provide services for companies to achieve successful customer data management. Lastly, an interview at DeLaval was conducted to view a company with similar issues and aims as the case company, which has implemented a central system for customer data management to achieve a unified view of their customers. Furthermore, telephone interviews were conducted with most interviewees few weeks after the initial interviews to ask further questions and clarify some points.

Scania

Scania’s objective is to deliver optimised heavy trucks and buses, engines and services, and thereby be the leading company in their industry. Scania operates in about 100 countries and has more than 35,500 employees. Of these, around 15,000 work with sales and services in Scania’s own subsidiaries around the world. About 12,300 people work at production units in seven countries and delivery centres in six emerging markets. Research and development operations are concentrated in Södertälje, Sweden, and employ some 2,900 people. Scania’s Head Office is located in Södertälje, the workplace for 5,300 employees. Local procurement offices in Poland, the Czech Republic, the United States, China and Russia supplement Scania’s corporate purchasing department in Södertälje. (Scania)

Traditionally, the Scania customers have had a direct contact with the retail dealers, who have developed the relationship, sold him products and services, and stored customer data in their own propriety database. Scania’s customers have evolved into multinational organisations that operate across borders. Today’s customers are complex and have a relationship with several entities within Scania. It is not only the home dealer, who maintains contact with the customer, but also national and foreign retail dealers, as well as corporate functions such as Scania Assistance and Fleet Management. The dealers and corporate functions each use their own database and work independently with their customer data management, which has resulted in a heterogenic management of customer data with a variance in content, structure and quality of the customer data. Scania aims to achieve a
unified view of their customers, though the sources of customer data are heterogeneous. (Boëthius)

**Ernst & Young**

Ernst & Young is a consulting firm that provide services within strategy, assurance, advisory, tax and transactions. (Ernst & Young) Interviews were conducted with Håkan Johansson and Muhammad Samad at the firm’s Stockholm office. Both senior consultants work within the firms IT Advisory, and mostly with information management projects. Johansson is responsible for Information Management at the IT Advisory department and has worked with Business Intelligence for many years. Samad’s experience lies in IT architecture and Customer Data Management within Financial Services. (Johansson; Samad)

**Xlent**

Xlent is a consulting firm specialised in Strategy, Business Integration and IT. Xlent provides services and IT solutions for management and maintenance of customer data. (Xlent) Lars Lindberg has been interviewed at Xlent’s Stockholm office. Lindberg is a senior consultant within Xlent’s IT department. He works with information architecture, and has created master data solutions for many companies within various industries. (Lindberg)

**DeLaval**

DeLaval has over 125 years of innovation and experience in the dairy business, supporting dairy farmers in managing their farms their way. DeLaval develops, manufactures and distributes equipment and complete systems for milk production. DeLaval has approximately 4500 employees and operates in more than 100 countries and caters to customers with livestock size ranging from 1 to 50,000 animals. (DeLaval) An interview has been conducted with Dan Oatley, the Business Development and Channels Support Manager at DeLaval. Dan Oatley works in the central marketing department and is responsible for the central CRM system, extranet and intranet. (Oatley) According to Oatley, DeLaval has a similar aim as Scania to achieve a unified view of their customers.

The benchmarking company, DeLaval, was chosen based on three criteria:
- Has several independent divisions that have contact with the customers
- Has multinational customers
- Has implemented a central system for customer data management

**Interviews**

Conducting interviews is a method used to collect qualitative data. Interviews can be done with individuals or groups, using face-to-face, telephone, email or videoconference methods. Under the positivist paradigm, interviews are structured, with questions that are planned in advance. Unstructured interviews, on the other hand, are when no questions are prepared before but evolve during the course of the interview. This type of interview is more commonly used for research conducted under interpretivism. The questions are open-ended with the purpose of exploring the interviewee’s opinions in depth. Unstructured interviews
are very time consuming and can be difficult to analyse. This type of interview is appropriate when it is necessary to understand the construct that the interviewee uses as a basis for his or her opinions, if one aim of the interview is to develop an understanding of the respondent’s “world” so that the interviewer might influence it, if the step-by-step logic of the situation is not clear, if the subject matter is highly confidential, or if the interviewee may be reluctant to be truthful about this issue other than confidentially in a one-to-one situation. (Collis, Hussey 2003)

The purpose of the interviews in the empirical study was to clarify how customer data can be managed successfully. Therefore, face-to-face interviews with semi-structured interviews were used. The questions were prepared in advance. Prepared questions are used to eliminate possible irrelevant information (Lantz 2007). The interview questions were created based on the findings of the literature study. A description of the topics that would be brought up during the interview, meaning no specific questions, was sent in advance to the interviewees. The reason was to enable reflection without influencing the interviewee beforehand. Interviews were recorded with a digital recorder to enable the interviewers to focus on managing the interview and notice non-verbal communication such as body language. Immediately after each interview, the recorded interview was transcribed. The transcription was then summarised and structured to easier identify relevant information. Interviews were conducted at the firms’ own offices. The duration of every interview was approximately one hour.

Interviews with Scania stakeholders were set up with the help of the thesis assignor, Daniel Boëthius, who is the Vice President of Franchise Standards and Volume Planning at Scania. Many Scania interviewees were contacted beforehand by Boëthius, while others were contacted by the writers of this thesis based on referrals made by other Scania interviewees and employees. Most interviews were conducted at Scania’s central office in Södertälje, Sweden with the exception of a few interviews conducted at the office of the interviewees also located in Södertälje. Interviewees from the consulting firms and benchmarking company were contacted by the authors of the thesis through telephone, by calling the telephone numbers displayed on the firms’ country websites. Interviews were conducted at the companies’ offices in Stockholm.

**Interviewees at Scania**

A description of the interviewees at Scania is represented in table 2.1.

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
<th>Department</th>
<th>Relevant experience for thesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daniel Boëthius</td>
<td>Vice President</td>
<td>Franchise Standards</td>
<td>Assignor of thesis</td>
</tr>
<tr>
<td>Mikaela Andersson</td>
<td>Franchise Communicator</td>
<td>Franchise Standards</td>
<td>Manager of the Scania International Service register (SIS)</td>
</tr>
<tr>
<td>Name</td>
<td>Title</td>
<td>Company</td>
<td>Task</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lars Grufman</td>
<td>Commercial Manager</td>
<td>Scania Assistance</td>
<td>Customer data management and invoice management</td>
</tr>
<tr>
<td>Michael Hedgren</td>
<td>IT Manager</td>
<td>Scania Assistance</td>
<td>In charge of Scania Assistance’s customer database SCUD</td>
</tr>
<tr>
<td>John Kernot</td>
<td>Information Security Architect</td>
<td>Scania Networks</td>
<td>Currently working on mapping customer information flow within Scania</td>
</tr>
<tr>
<td>Anders Bredenberg</td>
<td>Business Information Architect</td>
<td>Scania Networks</td>
<td>Currently working on mapping customer information flow within Scania</td>
</tr>
<tr>
<td>Lars Wiberg</td>
<td>System Analyst</td>
<td>Scania Networks</td>
<td>Currently working on mapping customer information flow within Scania</td>
</tr>
<tr>
<td>Anita Linder</td>
<td>Information Architect</td>
<td>Scania Networks</td>
<td>Currently working on mapping customer information flow within Scania</td>
</tr>
<tr>
<td>Martin Olsson</td>
<td>Information Resources Manager</td>
<td>Strategic IT Development</td>
<td>In charge of the Trading Partner Information project</td>
</tr>
<tr>
<td>Koen Knoops</td>
<td>Vice President</td>
<td>Financial Services Insurance</td>
<td>Involved in introducing a central data warehouse for the department</td>
</tr>
<tr>
<td>Lars Pålsson</td>
<td>Credit Manager</td>
<td>Scania-Bilar Sverige, Kungsens Kurva</td>
<td>Works with the development of Automaster, the dealers central database</td>
</tr>
<tr>
<td>Björn Winblad</td>
<td>Head of Scania Mining</td>
<td>Scania Mining</td>
<td>Leading Scania Mining through its first period as an independent department</td>
</tr>
<tr>
<td>Karin Rådström</td>
<td>Product Director</td>
<td>Scania Fleet Management</td>
<td>Head of department’s market side. Sells and performs quality assurance for services</td>
</tr>
<tr>
<td>Jesper Lovendal</td>
<td>Information Architect</td>
<td>Scania Fleet Management</td>
<td>In charge of the technical side of operations functions</td>
</tr>
<tr>
<td>Fredrik Goetzinger</td>
<td>Business Analyst</td>
<td>Scania Fleet Management</td>
<td>Knowledge about the different processes and services delivered</td>
</tr>
</tbody>
</table>

*Table 2.1: List of case companies interviewees*
External Interviews

A description of the interviewed consultants and the interviewee at DeLaval is represented in table 2.2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
<th>Company</th>
<th>Relevant experience for thesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Håkan Johansson</td>
<td>Information Management Consultant</td>
<td>Ernst &amp; Young</td>
<td>Senior consultant working with Information Management</td>
</tr>
<tr>
<td>Muhammad Samad</td>
<td>IT Architecture Consultant</td>
<td>Ernst &amp; Young</td>
<td>Senior consultant working with IT architecture and Master Data Management</td>
</tr>
<tr>
<td>Lars Lindberg</td>
<td>Information Architecture Consultant</td>
<td>Xlent</td>
<td>Senior consultant working with information architecture and Master Data solutions</td>
</tr>
<tr>
<td>Dan Oatley</td>
<td>Business Development and System Support Manager</td>
<td>DeLaval</td>
<td>Responsible for the CRM system, extranet and intranet</td>
</tr>
</tbody>
</table>

Table 2.2: List of external interviewees

2.5 Validity, Reliability and Generalisability

Validity is the extent to which the research findings accurately reflect the studied phenomena (Collis, Hussey 2009). Well accepted qualitative methods were used for the empirical study such as semi-structured interviews, which are characteristic for interpretivistic studies. Interpretivism focuses on capturing the essence of the phenomena and extracting data that provide rich and detailed explanations (Collis, Hussey 2009). The aim of such a research is to gain access to the knowledge of those involved in the studied phenomenon, which consequently results in high validity under an interpretivistic paradigm (Collis, Hussey 2009).

On-going communication with both assignor and supervisor has been held to ensure that the thesis progress is in the right direction to increase the validity. There is risk for misinterpretation during interviews, which affects the empirical results. Therefore, the material generated by the empirical study was sent to the interviewees for confirmation and approval, to confirm the accuracy of the empirical results and thereby increase the validity of the thesis.

Since the interviews only present the opinion of one person, the authors have made sure to conduct as many interviews as the time frame has allowed. A survey could have been used to complement and confirm the empirical results to generate a more reliable result, giving a broader view of the issues. A survey was however not chosen due to the limited time frame.

This study can be applied to Scania as it is the main case study. However, it can also be interesting for organisations that implement or plan to implement Master Data Management and that collect data from heterogeneous sources. The empirical study
involved a limited number of Scania divisions, a benchmark company and two consultancies. One can therefore not argue that the results are universal. However, by including interviews of two major consultancy firms on can argue that the empirical results give a good indication of real world issues and factors playing a role in the implementation of Master Data Management.
3. Literature Study

The literature study covers the existing knowledge and literature regarding the topic Customer Data Management (CDM) within the concept Master Data Management (MDM). Firstly, the crucial factors and challenges regarding CDM, according to the literature are presented. Since the issues revealed by the literature stress organisational challenges regarding achieving acceptance of changes, theories concerning Change Management are included. Further, the fundamentals of CDM relevant for and within the scope of the thesis are presented and are the following: Data Governance, Data Stewardship, Data Quality Management, and Data Quality Assessment and Maintenance. The literature regarding Data Quality Management and Data Quality Assessment and Maintenance includes theories both within and outside of the concept of CDM to get a broader perspective.

3.1 Customer Data Management

MDM can be defined as a set of procedures and technologies for collecting, and maintaining high quality master data. Master data is data that is used by several business units of the enterprise. MDM involves more than applications and technologies, it also requires an organisation to implement policies and procedures for controlling how master data is created and used. (White 2007) MDM has a management focus and not a technology focus. Introduction of new technologies without data management thinking does not result in a unified information system. (Dayton 2007) Although MDM should not be considered a technology project, it cannot be done without using tools and technology to support the initiative. (Loshin 2009)

One of the main objectives of an MDM system is to publish an integrated, accurate, and consistent set of master data for use by stakeholders (White 2007). With proper governance, the master data can be regarded as a unified set of data that all applications can rely on for consistent, high quality information (Loshin 2009). The integrated set of master data is the golden copy and is the single place in an organisation where the data is guaranteed to be accurate and up to date (White 2007). According to Loshin (2009) successful MDM relies on the following:

- Methods for identification of master data
- Unified definitions of data across business units
- A high quality integration technology
- A governance framework for managing continued integration of enterprise data into the master data environment

One key challenge when managing customer data is to collect relevant and accurate data. Furthermore, the accuracy of the data needs to be maintained in all the systems the data is included, despite changes made (Dayton 2007). Additional key issues are unclear and vague process definitions for collecting and maintaining data, and lack of clarity in data ownership. Furthermore, maintenance of data is challenging because of the complex and continuously
increasing amount of data. (Silvola et al. 2011) According to Loshin (2009) the greatest challenges for integrating data from multiple sources into one master data are organisational rather than technological.

With proper CDM, firms can achieve a “single version of the truth” about their customers by collecting and maintaining clear, accurate and centralised customer data (Dyché, Levy 2006).

3.2 Crucial Factors and Challenges

The essential factors and common challenges involving the introduction of a CDM initiative identified by Loshin (2009) are presented:

- **Acceptance of Change:** The introduction of a CDM program leads to changes in how business processes are defined and executed. Therefore, it is difficult to convince all the employees that are affected by the changes to accept the introduction. (Berson, Dubov 2007) Training and individual incentives must be in place in order to create a smooth transition from vertical information sharing to a collaborative one. Further, it is crucial that managers clearly communicate the benefits of the change. (Loshin 2009) This challenge is the most difficult one when implementing a CDM initiative according to Berson and Dubov (2007), and Loshin (2009).

- **Reaching Consensus:** The introduction of the use of master data requires that all stakeholders reach consensus on work procedures, share their data, and provide efforts for data quality improvement to ensure that the result meets the data quality requirements of all organisational entities. Integrating information from the entities of a company into a single source of information implies that the entities of the business adapt to the needs of the enterprise. (Loshin 2009) Reaching consensus is difficult because of the complexity of stakeholder’s landscape. CDM efforts involve many stakeholder groups; including executive management, line managers, front office, back office, data governance experts, technology managers, information architects, finance, legal department, information security, and sales. Their difference in priorities and objectives connected to the different business processes complicates the developing of a unified strategy and objectives for the CDM efforts. (Berson, Dubov 2007) It is important to define clear policies and procedures for governing the maintenance of the master data. By seeking comments from across the organisation, it is possible to create a framework through consensus (Loshin 2009). Thereby, CDM efforts focus to a large degree on political and consensus-building activities. (Smith, McKeen 2008)

- **Unified Definition for Data Elements:** Data that is collected in isolation will have a variety of definitions, formats and representation. For example, a customer according to the sales department is any prospect, while the accounting department defines customer as a party that has agreed to buy their product. This becomes a problem when integrating the business units’ data into a single view master data.
It is difficult to have all stakeholders agree on common definitions for the data used in their business (Smith, McKeen 2008). Loshin (2009) suggests that the solution is to establish guidelines for the definition of data elements by conducting workshops with stakeholders.

- **Integration**: The challenge is to be able to quickly and accurately capture, standardise, and consolidate the large amount of customer data that comes from a variety of channels, entry points, and systems. The problem most companies face is the inability to accumulate a complete customer view when most of the systems are isolated from each other and operate independently (Berson, Dubov 2007). Further, companies make the mistake of centring the business processes around a chosen technology, when in reality technology should be put in place to support the business processes instead of processes being developed around the technology (Loshin 2009).

- **Data Stewardship**: The business units of most firms can be considered as islands of information due to years of vertical information sharing. Therefore, it is difficult to assign responsibilities in relation to the isolated data sets as they are integrated into a master data system. Customer data management transfers the responsibility and accountability for information management from the business units to the organisation. The benefits are that one individual or group is responsible for the single version of accurate data, instead of several individuals being responsible for different versions of the same data objects. However, the challenge lies in how the stakeholders react to the reassignment of ownership. (Loshin 2009) One key challenge is to identify the owner of each piece of data (Smith, McKeen 2008). Smith, McKeen (2008) believes that it is likely that co-ownership is needed.

### 3.3 Change Management

Since the challenge concerning acceptance of change is considered the most difficult one, theories within Change Management are used for comparison, which are concerned with challenges when introducing any new project or initiative. Kotter (1996) has developed an eight-step model, which explains how change can be lead and managed successfully. The eight-step model is presented in figure 3.1.
According to Kotter (1996), the challenges of introducing change are connected to these eight steps.

1. **Establishing a Sense of Urgency**: Creating interest and commitment from employees is one of the most important factors, and also the most difficult one, when implementing change. It is essential that commitment is established by identifying and discussing potential crisis or lost opportunities that can occur if change is not introduced.

2. **Creating a Guiding Coalition**: It is important to form a group that has the authorisation and competence required to lead the change. Companies often underestimate the difficulties of producing change and therefore the importance of this step.

3. **Developing a Vision and Strategy**: A clear and convincing vision is developed. The vision describes the objective and its purpose is to direct the change efforts and create interest for the change project. The challenge is to develop a vision that is clear and comprehensible for all employees involved. After this, strategies for achieving the vision are developed.

4. **Communicating the Change Vision**: The vision is then communicated to all employees that are affected by the change, in a repetitive manner. Transformation is impossible if not most employees are involved in the change efforts.

5. **Empowering Employees**: Possible obstacles are the organisational structure, narrow job descriptions, compensation or performance-appraisal systems, or lack of leadership directed to the change efforts. Structures and processes that counteract the change efforts must be
replaced with ones that gives authorisation and encourages acting towards the objectives of the vision.

6. **Generating Short-Term Wins**: Initially, short-term positive results are actively created to show the gains acquired by the change project and in this way, create interest and commitment by the employees. Positive results in the direction of the change efforts must be recognised and the employees involved be rewarded.

7. **Consolidating Gains and Producing More Change**: The challenge at this stage is to not declare victory too early as soon as positive results are recognised. The gained trust from the short-term wins is used to further eliminate obstacles and produce more changes.

8. **Anchoring New Approaches in the Culture**: The last step is to integrate the changes made into the corporate culture to ensure their continued existence. The integration is achieved by communicating clear connections between the changes made and the corporate success. This last step is often neglected since the efforts can be assumed to be finished when the changes and improvements are implemented. Though, change is retained only when it becomes part of the corporate culture.

Commitment and leadership from management is essential for change to be successful since all of the steps above are dependant of managerial commitment. Management on every level, from CEO to middle-level executives, must be involved and committed to the change efforts. “A majority of employees, perhaps 75 percent of management overall, and virtually all of the top executives need to believe that considerable change is absolutely essential.” Managers are responsible for introducing policies and an organisational structure that supports the change efforts. Further, managers on every level must communicate the benefits of the change, and the consequence of not working towards the change. (Kotter 1996)

### 3.4 Fundamentals of Customer Data Management

Based on the crucial factors and challenges, focus is given to the fundamentals of CDM that are related to the issues identified: Data Governance, Data Stewardship, Data Quality Management, and Data Quality Assessment and Improvement.

**Data Governance**

Data governance is one key success factor when introducing MDM (Dreilbelbis et al. 2008); data governance is the glue in MDM (Cervo, Allen 2011). In a study conducted by The Information Difference, survey findings show that 88 percent of the respondents, who mostly come from large companies in North America and Europe, felt that implementing data governance either prior to or together with a MDM initiative is critical to the success of the MDM efforts (The Information Difference Company 2010). Data governance involves effectively using people, processes, and technology to leverage the master data as an asset (Dreilbelbis et al. 2008). The purpose of data governance is to ensure that the data meets the expectations of all the business objectives (Loshin 2009). The objective of data
governance is to create a framework for continuous data quality measurement and improvement, by introducing policies, processes, activities, and an organisational structure that benefits MDM objectives (Berson, Dubov 2007).

Cervo and Allen (2011) suggest that data governance entails two phases: planning and implementation. These are represented in figure 3.2.

Figure 3.2: Phases of MDM Planning and Implementation based on the description of governance by Cervo and Allen (2011)

One of the first and most serious mistakes commonly made when implementing a data governance plan is that a data governance plan is not considered early enough. Data governance needs to be clearly distinguished from other types of governing organisations that typically exist in a company. The objectives and the scope of the initiative are defined. Further, clear roles and responsibilities for data maintenance are defined. The roles and responsibilities are not necessarily connected to specific titles or levels. (Cervo, Allen 2011)

The value of customer data is highly dependent on the accuracy of the data. An important factor for accurate data depends on data entry processes. According to Loshin (2009), information policies are defined to enable development of definitions of the data quality requirements associated with each data element. The data quality requirements are then used as validation rules for data quality control. When the data governance plan and requirements are set, an improvement of data quality is supposed to be shown. To do this, there needs to be information available about the current state of the data quality. (Cervo, Allen 2011) Loshin (2009) suggests that the information architecture must be mapped since databases often are created in isolation and formed to support functional requirements for a specific business unit, without considering any overlaps with other business applications. Before a data governance framework can be introduced, management must understand, and document the actual information architecture. The documentation involves what data assets exist, how the data is managed and used, and how they support current business. (Loshin 2009) Overlaps of data used in different databases are detected by evaluating the information architecture. Further, data monitoring processes are developed to be able to control and improve data quality according to the requirements defined. (Cervo, Allen 2011)
Cervo and Allen (2011) continue the data governance model into the implementation phase. Further, the company must validate the defined processes to ensure that they are ready for implementation. The next step includes implementing the concept. For the implementation to be successful, awareness regarding data quality must be maintained by continuously informing stakeholders about the on-going efforts and their benefits, and keeping the employees given the data governance support roles engaged. Regular meetings should be held with the sole purpose of keeping awareness and communicate activities and the connected benefits. (Cervo, Allen 2011) The remaining, and essential, step is to maintain a steady state and generate further data quality improvements. On-going data quality and process evaluation and training are necessary to maintain a high quality management of customer data. Data entry is monitored to detect and correct negative data entry behaviour; processes are regularly revised and improved. Furthermore, the achieved benefits are communicated since success breed’s opportunity for further success. (Cervo, Allen 2011)

Data Stewardship

One of the largest challenges with data governance is the lack of follow-through. Well-defined governance policies do not lead to successful data governance unless the underlying organisational structure is in place. Clear roles and responsibilities must be introduced for efficient data governance. These roles are called stewards. However, the governance framework supports the needs of all stakeholders, and therefore benefits from participation from across the enterprise. (Loshin 2009) It is important to acknowledge that the “who”, “where” and “how” of data stewardship is highly dependent on how the overall CDM practices are defined and executed (Cervo, Allen 2011).

Maintenance of data is executed to a great deal by the stewards. All data must have a steward who is responsible for ensuring the quality of the data. The steward is usually an individual that has knowledge of the business and can recognise incorrect data, and has the authorisation and knowledge to correct the issue. The steward is also responsible for maintaining a regular relationship between the data creators and users (Cervo, Allen 2011). Their role is to continuously measure and assess data quality to ensure that the data quality requirements are met (Berson, Dubov 2007).

Clear roles and responsibilities are crucial for successful CDM. However, it is important to note that an objective is to create a culture of focus on data quality. Therefore, processes that enable and encourage data quality improvement should be available for all employees. (Cervo, Allen 2011)

Framework for Responsibility

One of the principles of MDM is that master data is an enterprise asset. Therefore, the ownership of the master data is given to data stewards on a central level. The central data stewards oversee the master data: defining the rules and policies relating to the master data, ensuring that the master data is properly controlled, that the master data quality is good, and promoting the MDM efforts across the organisation. (Dreilbelbis et al. 2008)
According to the literature found on the subject of data stewardship, there should be several levels of customer data responsibility in a company (Karel; Loshin; Cervo, Allen). Loshin (2009) suggests that the roles that should be introduced are the following:

- **Data Governance Director**: The data governance director is responsible for the everyday data governance. The director administrates guidance to stakeholders and oversees the use of the information policies. The director chooses the data governance oversight board. The need for governance efforts and reporting of the data governance performance are part of the data governance’s responsibilities.

- **Data Governance Oversight Board**: This board is composed of representatives from across the enterprise, and their responsibility is to lead and oversee data governance activities. The governance oversight board provide strategic direction for data governance, review information policies and assign groups to define information policies and rules based on business policies, approve data governance policies and processes and manage the reward frameworks.

- **Data Coordination Council**: The data coordination council is responsible for the management of the actual governance activities; they oversee the work of the stewards. The council consists of a group of stakeholders from across the company. It is the council’s responsibility to adjust processes continually to ensure that the data quality requirements are met.

- **Data Steward**: The data steward is responsible for continuously evaluating data, handle problems with data and ensure that information policies are followed. It is the steward’s task to communicate issues to the stakeholders affected.

Loshin (2011) suggests that the data stewards are responsible for:

- Supporting the data users by collecting and prioritising data issues. The issues communicated to those that are affected, and either resolve or communicate the issue to those who can resolve it
- Maintaining data by regular updates, and ensuring that the resources required for updating data are available and working properly
- Overseeing data quality by defining the data quality rules, and assess and improve data quality. The steward must oversee the quality of data, communicate changed business requirements and participate in the implementation of data quality standards
- Validating data as it enters to ensure the quality of the data.
- Distribute information regarding customer data management to stakeholders throughout the enterprise

Karel (2007) defines four key data governance roles:

- **Executive sponsor**: Is appointed at a high executive level to increase the potential for enterprise adoption. This role should be identified early and is responsible for driving the efforts forward.
• **Program driver:** This individual or team is responsible for communication between stewards and executive sponsor, coordination of stewards, and on-going auditing of data quality and metrics.

• **Stewards:** Data stewards are divided into business stewards and IT stewards. The stewards are the customer data experts with focus on business and IT respectively. They are responsible for education and support within their area of expertise towards the stakeholders.

It is important to note that these are roles and not titles. The governance responsibilities might be only a part of the individuals overall responsibilities. (Karel 2007) Data stewardship is not necessarily an information technology function, nor should it necessarily be a full-time position. Data stewardship is a role that has a set of responsibilities along with accountability to the line-of-business management. (Loshin 2011)

Cervo and Allen (2011) propose that roles and responsibilities should exist on two levels. The data stewards on the first level focus primarily on business support and data quality management of the master data. Cervo and Allen (2011) suggest that the data stewardship model should also include data stewards representing the different business units where primary creators, updaters and data users are.

The stewards on the first level are considered the owners of the overall CDM process, with a focus across various roles and business functions. The main responsibilities of the central stewards are:

• Coordinating and managing the CDM processes
• Leading the development of data quality requirements and metrics by measuring and monitoring performance, and improving data quality, and supporting other employees in data quality improvement efforts
• Constantly communicating issues regarding CDM to higher level executives and other stakeholders
• Engaging in IT activities that support CDM (Cervo, Allen 2011)

The purpose of the stewards representing the different business units are to maintain the CDM processes and high data quality in a coordinated fashion across the business units. These data stewards work closely together with the central stewards, and are responsible for every day CDM efforts. The stewards need to have two types of expertise; business processes and CDM concept and practice. The main responsibilities of the business unit stewards are:

• Managing CDM initiatives on a local level
• Enforcing CDM policies and standards locally. The data stewards should be given adequate training to understand how and where policies need to be applied
• Providing feedback regarding the development and implementation of policies and standards
• Monitoring and controlling data quality. The data stewards need to coordinate with the central stewards to determine what specific areas of data quality are essential to monitor and control.
• Managing the incident management process, by raising and help resolve issues. (Cervo, Allen 2011)

Data Quality Management
Data quality can be defined differently depending on the stakeholder. The definition of data quality is often different between the consumers of data and the creators of data. (Wang, Strong 1996) The level of the perceived data quality is highly subjective and dependent on who uses the data and the purpose intended for the data. Data whose quality makes it highly useful for one application could be of insufficient quality for another application to function. Applications have different requirements on the quality of the data in order to function as expected. (Batini, Scannapieca 2006). Data requirements are formulated into data dimensions to better assess data.

Data Dimensions
Data quality dimensions are sets of data quality attributes (Wang, Strong 1996). The definition of the dimensions and metrics to assess data is a critical (Batini, Scannapieca 2006). Studies have shown that several dimensions are needed to define data quality (Wang et al. 1992). According to Loshin (2009) data quality expectations are defined by stating dimensions. Data quality dimensions are then used in data validation and monitoring to ensure that expectations are met (Loshin 2009).

There is no widely accepted agreement on which set of dimensions that are more appropriate to define data quality or how the dimensions should be defined (Batini et.al 2009). Batini et al. (2009) conclude, after analysing several authors’ suggested dimensions, that there is a set of dimensions that most authors have in common. Their analysis has been compared to Loshin’s (2009), Pipino et al. (2002) and Wand and Wand’s (1996) opinion on the matter, and the following common dimensions have been identified:

• **Accuracy:** According to Batini et al. (2009) and Loshin (2009), accuracy is the extent to which data are correct, reliable and are compatible with the values of the real-world entities it represents. According to Batini et al. (2009) accuracy can be distinguished into semantic and synthetic accuracy. Semantic accuracy is a measure on the proximity of a data value to a value that is considered to be correct. Synthetic accuracy measures how close a value is to a set of values. For example, “Jean” is semantically close to “John”. A reference data set is needed in order to measure accuracy with an automated process; otherwise a manual process is needed in which the provider of the data must be contacted to confirm the accuracy of the data if the data is collected from another source. Accuracy is a challenging dimension to monitor
since the real world information that the data is supposed to represent changes over time. (Loshin 2009)

- **Completeness**: Completeness is the degree to which a given data collection includes data that describes real-world objects. In other words, how many missing values that exist in a given data collection (Batini et al. 2009). Completeness can be identified in one of three ways. The first one is by a mandatory value assignment, meaning the data element must contain a value (Loshin 2009). The missing values can be a result of data that is known but is unavailable or just simply does not exist (Batini et al. 2009). The second way of identifying completeness involves forcing data element to have or not to have a value under certain conditions. The third way of identifying completeness is about which data element values are applicable. For example entering “waist-size” for a hat. (Loshin 2009)

- **Consistency**: This dimension refers to the degree of violation of semantic rules in a set of data items. Semantic rules regard the range of values that should be used for a data element and the elements that should to be used in a set of data (Batini et al. 2009). Pipino et al. (2002) definition of consistent representation, the extent to which data is presented in the same format, is suitable to describe consistency. Simply put two data values taken from two different data sets should not conflict with one another (Loshin 2009).

- **Timeliness**: According to Batini et al. (2009) time related dimensions are an important aspect of data. The literature study has shown that time related dimensions cannot only have different definitions but also different names. The main time related dimensions usually go under the name currency, volatility and timeliness. For this report, the definition formulated by Wand and Wand (1996) is most relevant as it includes both a subjective and objective evaluation, and is the most general one reviewed by authors of this report. Wand and Wand (1996) defines timeliness as “a measure of the extent to which the age of the information is appropriate for its purpose”.

- **Uniqueness**: Loshin (2009) suggests an additional dimension that is suitable in a MDM environment. Uniqueness refers to the existence of unique data within a data set and is characterised by the fact that no entity exists more than once. Data instances should then not be created if such instances already exist. This dimension can be evaluated through duplicate detection. (Loshin 2009)

**Data Quality Assessment and Improvement**

“There is no one-size-fits-all model for data quality” (Cervo, Allen 2011). It is necessary to take into account factors such as the corporate culture, the MDM approach being implemented, maturity of governance and stewardship, level of management engagement, personnel skills and technology resources (Cervo, Allen 2011).

When introducing the assessment processes, a light is shed on a large amount of data quality and data management issues. This is because of the previous lack of processes to
channel the existing issues. Therefore, the value of the efforts must be communicated to all customer data stakeholders. (Cervo, Allen 2011)

Companies must take into account both the subjective perception of the individuals involved with the data and the objective measurements made of the data. Subjective data quality assessments reflect the needs and experiences of stakeholders: data collectors, data stewards and data consumers (Pipino et al. 2002). Emphasising on the “fitness for use” aspect of data can lead to the wrong assumption that objective assessment is not possible (Batini, Scannapieca 2006). Objective assessment involves metrics that reflect states of the data without knowledge about the application, meaning that the metrics can be applied to any data set regardless the task involved, and metrics that take into account the organisation’s business rules, company government regulations, and constraints provided by the database administrator. (Pipino et al. 2002) Batini and Scannapieca (2006) believe that most data quality dimensions should have objective measures given that the perceived quality can be evaluated in relation to a given applications requirements, meaning that there would need to be a suiting set of objective measures for each application. The quality dimensions should be measured according to the given applications requirement (Batini, Scannapieca 2006).

Data Quality Assessment

There are several different methods for the assessment and improvement of data quality assessment. Batini et al. (2009) presents the steps and phases that are common for most assessment and improvement methods. These are divided into three major phases presented in figure 3.3.
Figure 3.3: Major phases for Data Quality Assessment and Improvement Methods based on Batini et al. (2009), Batini and Scannapieca (2006).

The steps of the assessment phase are (see figure 3.4):

1. **Data analysis**, is a step for examination of data schemes and performing interviews to achieve complete understanding of data, and related architectural and management rules (Batini et al. 2009). By interviewing data users one can better understand the negative effects of poor data quality has on activities performed by the users (Batini, Scannapieca 2006).

2. **Data quality requirements analysis** includes surveying the data users and administrators about their opinions in order to identify quality issues and set new quality targets (Batini et al. 2009).

3. **Comparing** the results of the assessment. Pipino et al. (2002) suggests that the objective and subjective assessment are to be compared in order to identify differences and determining the root causes of the difference.

4. **Identify the critical areas**, which involves selecting the most relevant databases and data flows to be assessed quantitatively. (Batini et al. 2009)

5. **Create Process**, a model of the process producing or updating data is created.

6. **Measure quality**, in this last step, the quality dimensions that are affected by the quality issues that were identified in the second step (Data quality requirements analysis) are selected and corresponding metrics are defined. Measurements can be
objective when based on quantitative metrics and subjective when it is based on qualitative evaluations. (Batini et al. 2009)

Figure 3.4: Steps of Assessment based on Batini et al. (2009); Batini, Scannapieca (2006); Pipino et al. (2002).

Data Quality Improvement

A data quality activity is “any process performed directly on data to improve quality” (Batini, Scannapieca 2006). Data quality activities or initiatives for improvement can be both reactive, meaning measures are taken when a problem arises; and proactive, where measures are taken in before hand to prevent or minimise the risk of problems. Drivers, people pushing for customer data management improvements, within the company will either be reacting to data quality issues or working proactively to avoid them. Users that encounter a data anomaly or corrupt data in their daily work will not necessarily know how to solve the problem or what the source of it is. There must be a routine for how to proceed; a mechanism for presenting the problem and asking for correction. (Cervo, Allen 2011)

After assessment, data quality is improved in the improvement phase, which includes following the steps (Batini et al. 2009):

1. **Evaluation of costs**: Estimation of direct and indirect costs of data quality.
2. **Assigning the process responsibilities**: Identifying the process owners and define their responsibility.
3. **Identifying the sources of errors**: Finding out the causes the errors.
4. **Selecting strategies and techniques**: Identifying all the data improvement strategies and corresponding techniques that comply with knowledge of the current data quality situation, data quality objectives and budget constraints.
5. **Designing the data improvement solutions**: Involves selecting the most efficient strategy and related set of techniques and tools in order to improve data quality.
6. **Process control**: Data quality should be monitored during the data production process by setting up checkpoints in the process.
7. **Process redesign**: Deciding on which actions are to be taken to improve the process in order to deliver the corresponding data quality improvements.
8. **Improvement management**: Defining new organisational rules for data quality.
9. **Improvement monitoring**: Establishing periodic monitoring activities that provide feedback on the results of the improvement process and enables its dynamic tuning.

Different methodologies for improvement use a number of different strategies for the improvement phase. Each strategy has a number of techniques associated with it that are not discussed in this report. These techniques involve record linkage, acquisition of higher
quality data, process control and others. Methodologies adapt two types of strategies in general:

- **Data-driven strategies** that focus on improving the data quality by directly modifying the value of data.
- **Process-driven strategies** improve data quality through redesign of the data creation and modification process. (Batini et al. 2009)

**Standardisation**

Data standardisation is about conforming data elements according to pre-defined rules and policies (Cervo, Dalton 2011). Standardisation is critical in order to achieve one value for each data element. Not having a standardised data structure can lead to different representations of the same object or element. (Dreibelbis et. al 2008) Data objects may have different structures and different names or be modelled and stored differently. For example, some applications only have one element for a customer’s name while another has split the name into two such as “First Name” and “Surname”. While these two applications may store the same content conceptually, they present it differently. Slight differences such as these prevent computer systems from identifying similarities between record instances. Even different naming or data types of data elements can confuse applications. (Loshin 2009) For example, “CUST_NUM” can be used as a name for the data element storing customer number in one application but “CUSTOMER_NUMBER” could be used for another (Loshin 2009). Misspellings and nicknames can also cause problems for when using a certain key to find the information (Dreibelbis et. al 2008). The same information might also be stored under the same name but represented differently, for example, as a numeric field instead of alphanumeric. (Loshin 2009)

According to Cervo and Dalton (2011), the benefits of data standardisation are:

- It becomes easier to validate data if there is a set of rules already in place for a given data value that can be used for testing of data validating.
- It facilitates the identification of duplicates as data is more easily compared.
- It facilitates the identification of root causes of errors as data becomes easier to compare with other data.

The benefits of standardisation show that it facilitates the implementation other activities such as data validation and duplicate detection. However, there are some challenges regarding standardisation. Besides technological and resource limitations, data governance becomes an important driver of proposing and collecting standardisation requirements. (Cervo, Allen 2011)
4. Empirical Study

This chapter begins with an overview of the case company, which is the base of the empirical study, for the reader to get better acquainted with the empirical setting. Furthermore, this chapter also consists of the contribution to the thesis from the empirical study done at Scania, two consulting firms and one benchmarking company. The interview questions were created based on the findings of the literature study. First, the situation today at Scania and DeLaval are presented. Secondly, the findings from the interviews conducted with consultants, who are experts within the subject, together with the interviews conducted at Scania and DeLaval are presented in a chosen structure that emphasises the essential issues that have been noted. Interviews from Scania, DeLaval and the consulting firms have been presented together to simplify comparison of similarities and differences. This chapter only contains references to the interviewees. This chapter will be followed by an analysis and discussion of the literature and empirical findings.

4.1 Empirical Setting

Scania is a leading manufacturer in the heavy buses and industry- and marine engines. Scania also sells and provides a large supply of service related products and financial services. Scania’s objective is to deliver optimised heavy trucks and buses, engines and services, and thereby be the leading company in their industry. Scania operates in about 100 countries and has more than 35,500 employees. Of these, around 15,000 work with sales and services in Scania’s own subsidiaries around the world. About 12,300 people work at production units in seven countries and delivery centres in six emerging markets. Research and development operations are concentrated in Södertälje, Sweden, and employ some 2,900 people. Scania’s Head Office is located in Södertälje, the workplace of a total of 5,300 people. Local procurement offices in Poland, the Czech Republic, the United States, China and Russia supplement Scania’s corporate purchasing department in Södertälje. (Scania)

“Scania’s identity is shaped after its customers and products: vehicles, services and financing, and by the people within the company, their values and work methods. The three core values are ‘Customer first’, ‘Respect for the individual’ and ‘Quality.’” (Scania)

Under the Board of Directors, the President and CEO have the overall responsibility for the Scania Group. The President and CEO jointly decide with the Executive Board on issues in its area of competency that are of long-term and strategic nature. These include the development of the company, marketing, pricing policy, capital expenditures and financing. The Executive Board also prepares such issues that shall be decided by the Board of Directors. (Scania Group Homepage)

Group Management consists of the Executive Board and are the heads of the corporate the five corporate units, Production and Logistics, Research and Development, Finance and Business Control, Sales and Services Management and Franchise and Factory Sales. The corporate units are responsible for carrying out the established strategies. Each corporate unit reports to one of the members of the Executive Board. (Scania Group Homepage)
Figure 4.1: Scania’s Organisational Structure based on information from Scania Group Homepage

Figure 4.2: Franchise & Factory Sales Organisation based on information from Boëthius
Figure 4.3: Sales & Services Management Organisation based on information from Boëthius

Interviews were conducted with stakeholders from different business units of Scania’s Sales & Service Management corporate and Franchise & Factory Sales corporate units (see figures 4.1, 4.2, 4.3) and the dealer within the Swedish market to better understand the situation today and the goals and requests for the future from the different departments.

Franchise Standards & Volume Planning

Scania meets its customers through more than thousands of dealers and service points, who each offer customers products and services. This product offering and the way it is marketed, delivered and supported build the Scania brand. The Scania organisation expects that the service given is provided in a standardised way. These expectations and requirements are described in Scania Franchise Standards. The department of Franchise Standards maintains, develops, communicates and audits dealers based on the present standards. The department is the owner of the Scania International Service registers (SIS), which is a central database that contains contact information to the entire Scania sales and service network. (Scania Inline B)

Scania Assistance

Scania Assistance is the service coordinator for Scania’s customers. A single telephone call can put the customer in touch with a professional service coordinator at any time. Their main responsibilities are to help truck; bus and engine owners find the nearest workshop and put the customer in touch with them in case of breakdown; to support the Scania workshops in their business; and to assist customers and workshops with cross-country payments. (Scania Inline C) The department has the most extensive customer database, SCUD, within Scania. (Hedgren)
Scania Networks

Scania Networks is responsible for supporting Scania departments with IT services that enables them to run their business. (Scania Inline A) The department’s architect group is currently working on mapping support applications, such as dealer systems, that support different entities in Scania and the information flow between these applications in order to find the next generation system solutions. (Kernot)

Strategic IT Development

The department works with concept development regarding Scania’s IT strategy, based on analysis of the situation in different markets, competition and IT trends. Currently the department is working on developing a central database called Scania Commercial Knowledgebase. (Olsson)

Scania Financial Services Insurance

Scania Financial Services Insurance is a Scania owned insurance company. The products offered by Scania Financial Services Insurance are financing for Scania’s customers buying Scania products and similar products from other sellers, and insurance for their bought products. (Knoop)

Scania-Bilar Sverige, Kungens Kurva

Scania-Bilar Sverige is the distributor of Scania trucks, busses and engines in the Swedish market. Their dealer network in Sweden consists of 24 dealers. Scania-Bilar Sverige, Kungens Kurva is the head dealer in Stockholm, Sweden. (Scania-Bilar Intranet)

Scania Mining

Scania Mining is a cross-functional department that is entirely focused on the mining industry and sells products that are used on mining sites all over the world. Scania Mining used to be part of the Construction segment at Scania but due to the fast growth of the mining industry, it has broken out and become an independent department. (Winblad)

Fleet Management

Fleet Management offers continuous monitoring of trucks owned by the customer via regular reports that provide an overview of each truck’s performance and location including carbon footprint of a transport operation (Rådström).

4.2 Current Customer Data Management at Scania

Scania has several independent entities with their own customer data. Every department has historically solved customer data storing independently resulting in customer data being stored in several parts in Scania and with a variety of content and quality. Today, customer data are collected separately in every market and in every entity. (Kernot)
Each department collects its data from the different markets through different channels, some from local branches but most of them from distributors. The distributors collect their customer information from the dealers in their market. Each entity stores the distributors’ customer data separately. Due to the fact that a customer can have contact with several departments, one customer can appear in various ways in every system, resulting in a multiple and inconsistent view of the customer. The problem is worsened by the fact that today’s customers are becoming more international and operating cross borders making it harder to define the customer. (Boëthius)

There is no clear and common definition of a customer at Scania, since customers can operate in more than one region or country and can consist of several different entities each of which having a relationship with different Scania departments. As a result, departments may see one customer represented as several different ones in their systems. There is a need to define a structure of relationships between these entities and organisations. Customer data in today’s business is very turbulent and there is a need to maintain the validity of the data. It is common to find out-dated, inconsistent and corrupt data in the systems due to lack of maintenance routines. (Boëthius)

Scania has in the past created national databases for dealers to create one customer ID for each customer in each country, which worked for a while. However, customers have since then become more multinational and have started to operate across borders. (Kernot).

Scania has realised that there are customer data that the different entities store that might be useful for other entities and departments. There is a need for common data throughout the enterprise; Scania has for some years tried to describe what common knowledge is and exists in the organisation. Common knowledge is data from different databases that are collected in one place. The aim at Scania is to provide a central golden copy with common data that can be used by all information systems and in which users can trust as being correct and up-to-date. Scania is also developing a central trade partner information database. (Kernot)

Scania Assistance, Scania Fleet Management, Scania Financial Services and Scania-Bilar Sverige all have a form of central customer database (Hedgren; Rådström; Knoops). Scania Mining however is a new department and has not come so far along with customer data management (Winblad). Every department faces similar challenges with customer data management and solve customer data related problems differently.

For Scania Assistance, the challenge is not the collection of data, but keeping track of the ownership of the vehicle and ensuring that the data is up to date. Since it is difficult for Scania Assistance to control this, they have to trust that the dealers and distributors customer information is accurate. (Hedgren) For Scania Fleet Management on the other hand, the main challenge lies in being able to change how the databases are connected.
(Lovendal) Today, the same customer data exists in many different databases, which Lovendal wants to see, be changed.

### 4.3 Current Customer Data Management at DeLaval

DeLaval used to have several systems for collecting and storing customer data. In the process of updating the systems, it was decided to use a central server for customer data that is synchronised with the local databases. However, different systems for customer data are still used around the world. Today, a CRM system is used. There are five markets that use this CRM system at the moment, which is a central database. Access rights are determined by the users’ location. The sales staff, service technicians and installation engineers use the system. There are two ways for the employees to access the CRM system. One is to log on to a web application, which is connected directly to the server. Any changes made via the web are made directly on the server. The second way to access the central database is through the local databases on the employees’ laptops. The local databases are used for input of customer data, managing stock and placing orders. The data is synchronised from the local databases into a central server, which then sends the data on to the Enterprise Resource Planning (ERP) system and to the other relevant local databases. (Oatley)

The Enterprise Resource Planning (ERP) system, which also stores customer data, is used in around 20 markets. However, the ERP system does not contain all the customer data that is needed by the sales and service department. The ERP system does store company name, address, credit limit, invoice, but not for example, who visited the customer last or what service contracts the customer has. This information is stored in the CRM system.

DeLaval has developed an extranet where dealers can log on to a website and create orders, and view some customer data. This extranet is used in 71 markets today, with 7 more planned. DeLaval is now in the process of planning the introduction of the CRM system to the 71 countries that currently use the corporate extranet system. This plan raises additional difficulties in markets where there are collaboration with dealers instead of direct sales around the world. (Oatley)

In the future, the dealers should be able to log on to get access to the CRM system as well and make updates via the extranet, so that the customer data can be shared with DeLaval and their dealers. The difficulty here is that some dealers might see this as a threat to their ownership of the customer. The purpose of the customer data sharing is for DeLaval to better understand their customers to enable development of better products and services. It is essential for the dealers to understand the benefits of using the CRM system. “One thing for sure is that dealers will only use the CRM tool if they see the benefits to their own business, so usability and making the benefits easy to see are essential.” (Oatley)
4.4 Customer Data Management

All interviewees agree on storing customer data in a central database, which is integrated with the local databases. Though, there are some differences on the solutions chosen for the centrally stored data.

The consultants at Ernst & Young explain that the awareness regarding the issue of customer data management is being more and more discussed within companies in most industries. The concept of MDM is one that is popular, though every company has to research and understand the current situation and business to be able to decide what concept is suitable for the managing of their customer data.

Olsson explains that customer information is collected separately in every market and in every entity. The desire is to instead store customer data on a central level. Since Scania’s customers are becoming more and more multinational, there is a need for sharing of customer information throughout the enterprise.

“If a customer buys a truck in Russia, and it breaks down in Spain, the customer expects the workshop in Spain to have all the customer information and history needed to provide good service” (Olsson). Scania Assistance, Scania Fleet Management and Scania Financial Services have all chosen to introduce a form of central customer database (Hedgren; Rådström; Knoops).

Scania Assistance stores customer data in a central database called Scania Assistance Customer Database (SCUD). The data stored comes from the distributors in all 48 markets where Scania Assistance is servicing. This makes SCUD the most extensive customer database within Scania. The distributors upload the customer data as text files into SCUD every night. The distributors collect the customer data from the dealers in their market. Each dealer uses its own local systems for storing customer data. (Hedgren) Thereby, the customer data stored in SCUD has different sources (Grufman).

Scania Fleet Management also has a central database, which collects customer data from SCUD to their own central database. (Rådström) Scania Fleet Management has considered using the Trade Partner Master (TPM), which is Scania’s central trading partner information store. TPM is being developed by Scania Networks. (Wiberg) However, even though Scania Fleet Management welcomes the initiative (Lovendal), TPM has not been chosen as a source to collect customer data from since it is still under progress and only contains customer data from four markets, compared to the 48 markets that SCUD collects data from. (Rådström) However, there have been discussions between Olsson, Bredenberg and Lovendal about potentially letting TPM and SCUD merge together with time. Scania Fleet Management also collects information regarding distributors, dealers and workshops around the world from Scania International Service (SIS). The Scania Fleet Management system does not have a two-way communication between the databases where data is collected, which only allows data for a customer to be imported once and never updated. Changes in a customer’s data are not updated in the Scania Fleet Management system, since the direct customer contact
is made through the dealers. The dealers are responsible for updating the customer data. Furthermore, changes in the central Scania Fleet Management database are not updated in the local dealer systems since the department sees the local databases as the place where the most updated customer data is located. The purpose of the customer data in the Scania Fleet Management system is to get a view of the existing Scania customers. (Rådström) The import of data of new customers is done every night to Scania Fleet Management’s central database (Lovendal). Instead of customer data having to flow through different databases, Scania Fleet Management would like the local dealer systems to send customer data to a central hub that is synchronised with Scania Fleet Management’s and other central databases. (Lovendal)

The local branches of Scania Mining around the world have their own customer databases. The information is not shared between the branches (Winblad). Scania Mining local branches are now starting to send a quarterly report, with a standard structure, on the potential mining customers that have been identified and sales per customer. This information will be the basis of a future customer database. (Winblad)

Scania Financial Services Insurance stores data centrally using a data warehouse where all customer data is uploaded from the systems of all local branches. Financial Services Insurances’ local branches operate in different markets and each one use their own customer databases. The customer information is not shared between the local databases. All customer information from the local branches is uploaded automatically to a central data warehouse on a monthly basis. (Knoops)

Scania dealers have their own systems. Most dealers have many systems, including one central system called Automaster. The Automaster collects data from the other dealers systems and other external databases. All employees of the dealer can access and make changes in the Automaster. The Automasters of different dealers are not integrated with each other. (Påhlsson) Scania owned workshops use standard dealer systems. Independent workshops and dealers however, have their own systems. In this case, the Scania owned distributor has a responsibility of keeping record of Scania customers of the independent dealers in its market. In the European markets, all distributors are Scania owned. (Hedgren)

The dealers complement the customer data in their local systems with information collected from Scania’s National Database Sweden, which contains all customers in Sweden and their service history. Dealers in Sweden also collect customer data from the Swedish Transport Agency’s vehicle registry. The vehicle registry contains all vehicle owners in Sweden. When a customer makes an order, an enquiry is sent from the dealer system Automaster to collect customer related information from the vehicle registry and Scania’s National Database Sweden. Collecting customer data from external sources is not possible in all markets. (Påhlsson)
Franchise Standards at Scania own the central database Scania International Service (SIS). SIS is a database containing contact information to the entire Scania sales and service network, which is all distributors, dealers and workshops throughout the world. Many entities within Scania, such as Scania Assistance and Scania Fleet Management, use SIS though it is not integrated with the local systems. The distributors have access to SIS with the rights to view and change the information regarding themselves, the dealers and workshops in the same market. Furthermore, any stakeholder at Scania can access the SIS database through a web service. (Andersson)

According to Samad at Ernst & Young, customer data should be stored in a central database, which is synchronised with the local systems of every entity. Lindberg at Xlent agrees and explains that it is not possible for all entities to use the same system for storing customer data, since the system used must meet the requirements of a specific business. Every entity defines a customer differently and it is natural that units use a system that fits this view. Since it is not feasible for a large company’s units to reach a unified definition of the term customer, solution is to keep the local systems and integrate the data into a central customer database. (Lindberg)

Lindberg describes two models for integration between systems in an enterprise: the river model (see figure 4.3) and the lake model (see figure 4.3). When comparing the two models, one realises that the river model has some limitations according to Lindberg. For example, system D has the highest probability of having incorrect information since the information has been integrated for the third time. This means that the data quality deteriorates with every system that is added to the model. Further, a problem arises if system C contains information that does consist in system A, but not in system B. “For example, if system A contains e-mail addresses, but not system B, the e-mail addresses will not be passed along to system B. Then this information that system C needs is lost.” In the lake model, the

---

**Figure 4.3:** River integration model (left) and lake integration model (right) based on the description given by Lindberg at Xlent
information that is shared by all systems is stored in a central system. The information is then integrated from the central system to the local ones. The benefit of this model is that all local systems get the same information from the central system. (Lindberg)

*The purpose of storing data centrally is to enable sharing of data between local branches of the same entity, or different units.*

Scania Financial Services Insurance has recently started storing their data in a central data warehouse which has not been launched for usage yet. The plan is for all local branches to be able to view Scania Financial Services Insurance’s customers in every market and their information. Furthermore, the information will be shared with Scania Assistance to provide them with customers’ insurance information (Knoops), which is highly valuable for Scania Assistance (Grufman). The sharing of data might mean that the monthly uploads need to be more frequent to fit Scania Assistance’s requirements for current data. (Knoops)

There is no information sharing about the customers’ activities at other dealers. Påhlsson would like an Automaster for all dealers in Sweden or even one for Europe since customers are becoming increasingly multinational.

DeLaval uses a central system for customer data that is synchronised with the local systems. This solution was chosen to enable sharing of data between employees of the sales and service department. Today, every meeting or contact with the customer is registered in a central database, which is shared with the employee establishing further contact. (Oatley)

*There is a difference of opinion regarding if the central database should contain the data that is shared by all entities or all customer data that is needed for the company’s entities.*

Financial Services collect all customer data from local branches into their central data warehouse. Since the data is collected from only one entity, all customer data that is essential for their business is collected. In the future, when they share their data with Scania Assistance, only the customer data elements that are interesting for Scania Assistance will be shared. (Knoops) Similarly, SCUD is adapted after Scania Assistance’s business and contains the data elements that are important for the department’s business. (Hedberg) Goetzinger agrees and explains that Scania Fleet Management must supplement the data collected from SCUD with other information that is needed e.g. customer e-mail address data, from another database. While these entities collect all of the customer data needed for their business, Olsson suggests that customer data should be divided into three different categories, namely core, data used by most entities; shared, data used by many and unique, data that is only used by a single or few entities. Only core and shared data is stored centrally. (Olsson)

Similarly, Lindberg explains that there are customer data elements that all entities in an organisation use and therefore can be shared. The shared data should be stored centrally within the enterprise. (Lindberg)
Currently, the architecture group at Scania Networks are in the process of mapping the information flow and structure within Scania. During the mapping process, it has been discovered that there are some overlapping information between the local systems used by the entities. (Bredenberg)

At DeLaval, the central database contains all the customer data that the entities need, though the employees using it have access to different views depending on geographical factors, and what role and responsibility the employer has. For example, sales employees want to see their customers and previous contact with them, while a quality person wants to look at all claims for the whole country or region. In order to define what customer data should be included in the central database, the business processes must be outlined. Then it is determined what customer data is needed to support the processes. (Oatley)

*The more entities and markets that use the central database, the more reliable and usable the customer data in it will be.*

Scania Fleet Management chooses to collect customer data from SCUD instead TPM because SCUD is used in more markets (Rådström), supporting the statement above. Further, Lindberg explains that, since the central system is updated with the information from the local systems, all systems contribute to the “truth” represented in the central system. Every entity views some data elements to be essential for their business and is therefore concerned with the quality of those elements. Because different entities are concerned with different data elements, this leads to a natural quality control of data, and together, they all contribute to the quality of the data in the central system, which means that the data quality is increased with the numbers of local systems involved in the system. (Lindberg) Oatley at DeLaval believes that the more the central system is used, the more focus is on improving the system and maintaining the data it contains. “The more people that jump on this idea, the more successful it will be” (Pålhlsson). Hedgren agrees, and states: “The more people that use the central system, the more the relating efforts will be prioritised within the company.”

**4.5 Crucial Factors and Challenges**

*One issue is lacking communication between entities and local branches regarding how to manage customer data. Entities act as isolated silos.*

Most entities within Scania have realised the importance of customer data management and have introduced projects for how to improve customer data management. Scania Assistance has the central database SCUD (Grufman), Scania Financial Services Insurance have started to collect data from all local databases to a central data warehouse (Knoops), Scania Networks is working on TPM (Kernot), and Scania Fleet Management have a central database where all their data is stored and their working on introducing a central hub where data is collected from local system (Rådström).
Implementation of a customer data management program is mostly a question of people and how they adapt to change, rather than technological difficulties. The challenges are typically organisational rather than technical. Many companies make the mistake of implementing the technical solutions first and then trying to build the organisation around it. The choice of technical solution must be based on the organisation and its business processes.

According to Goetzinger, what is important for Scania Fleet Management’s customer data management is that there is an organisation and processes in place. Lovendahl adds that one should not focus too much on the technology, but instead focus on the routines and organisation for customer data management. Similarly, Samad at Ernst & Young explains that what needs to be done is structuring the business requirements and streamlining the governance process based on the requirements. The requirements are defined by having workshops with all the entities, to understand the needs and requests. Then an IT solution is chosen to fit the requirements and processes that are in place. Flexible technology is needed to support the processes, though it is neither the most difficult challenge nor enough to achieve successful customer data management. Entities must adapt to each other and share processes and routines to enable integration between the databases. (Samad) The challenges regarding governance involve people. Processes and technology might be in place; however success is achieved only if the employees have the knowledge and motivation to follow the directions. The difficulty occurs because the employees of different entities find it problematic or unnecessary to adapt to new technology and processes. Therefore, clear managerial communication regarding the need for and benefits of the change is key. (Johansson)

Implementing the technology is one of the last steps of implementing a CRM system. Many processes must be in place before a technology solution can be implemented. “A lot of companies just implement the technology thinking that this will solve all of their problems. They make the mistake of implementing the technology and then trying to put everything into place after.” Before any technology is implemented, the business processes must be mapped and understood. “You need to understand how you do things today, why you do it in this way, if it is the best way to do it, and if you can do it better”. Based on this, the most efficient processes are chosen. Stakeholders must be identified and involved in this process from the beginning. They will be the advocates when the change is introduced, since they have been involved in influencing the change and the new processes. The biggest challenge has been, and is still, getting people to change the way they work. The key is to determine the benefits of the change and communicate this to all employees involved. (Oatley)

The definition of customer is based on the business. It is therefore not possible to have a unified definition of a customer, throughout an enterprise with several business units. Though, it is believed that it is possible to develop unified definitions on customer data elements.
According to Winblad, a customer is someone that has the majority of its fleet or operation in a mining site. Grufman means that a customer is defined as the owner of a Scania vehicle or engine. An exception is when the owner is a leasing company, in which case the user of the vehicle is considered as the customer. (Grufman) However, Hedgren explains that the owner of a vehicle or engine is registered as the customer in SCUD, since there are no incentives or demands for leasing companies to inform Scania Assistance about which are the users of the vehicle or engine. According to Pålsson, both owner and the user of vehicles are represented in their central database, Automaster. According to Knoops, a customer is primarily a company that has insurance or financing through Scania Financial Services. Furthermore, future customers are also viewed as customers. Future customer is defined as a Scania customer that a dealer is in the process of selling insurance to. (Knoops)

Olsson believes that it is impossible to reach consensus throughout the organisation on the context and function of the customer data elements, since the definition of customer varies between entities. Lindberg believes that every entity defines the term customer differently. However, Olsson adds that it is important to reach consensus on a unified definition on customer data elements, even though the definition of a customer is different between entities. He suggests the OAGIS standard is being used for this purpose. OAGIS is an international standard that defines key naming conventions and key common content of data elements. Using the standard throughout an enterprise will lead to all business languages to be based on the same concepts. (Olsson)

*Clear ownership roles must be in place to achieve successful customer data management and thereby high data quality.*

Grufman firmly believes that the data quality depends on the customer data management on a dealer level. “The key to solving data quality issues is to have clearly defined data ownership and clear areas of responsibilities.” (Grufman) Hedgren believes there also should be a central organisation within Scania responsible for customer data.

For the maintenance of high data quality, it is crucial to have roles that are responsible for customer data; otherwise, if no one considers the maintenance of customer data quality as a part of their job, the work will not get done. (Samad) “It is important to know who does what” (Johansson). Clear ownership and custodianship of customer data is essential. (Johansson)

There needs to be a customer data custodian, who is responsible for customer data, on a central level. If there is no clear role for monitoring data, the data will not be monitored since it is not included in any other job description. Data should be updated on an entity level since all employees should be able to make changes in their own system. When changes are made, some data elements can be updated in the central system immediately while other elements need to be approved by the custodian before the data is spread. (Lindberg)
One essential factor is to decouple customer ownership and customer data stewardship.

Sales employees are protective of their customers. Therefore, it is important to decouple stewardship of customer data, and ownership of customer. “It is important that, for example, the sales employee understands that even though the customer data might be owned by another role, the ownership of the customer is not lost”. (Samad)

DeLaval has developed an extranet where dealers can log on to a website and create orders and view some customer data. This extranet is used in 51 markets today. In the future, the dealers should be able to log on to the central CRM system as well, so that the customer data can be shared between DeLaval and their dealers. The difficulty here is that some dealers might see this as a threat; “They might think, ‘You want to take our customers away from us and then you will not have a need for us anymore and our business might be taken away from us as well’”. Though, the purpose of customer data sharing is for DeLaval to better understand their customers, to enable development of better products and services rather than to take ownership of the dealers’ customers. (Oatley)

One issue regarding managing of customer data and maintaining high data quality is data entry errors.

The different entities have identified different sources of errors but in general all of them consider input errors to be a cause of error. At Scania Financial Services Insurance the danger lies within the registry of the customer’s assets, which are the equipments installed on the chassis. The asset related data does not come from production, which is why the data has to be entered manually into the systems.

Scania Assistance is well aware of what situations can cause data anomalies. For Scania Assistance data anomalies can arise when the shift of data ownership has not been made clear when vehicles are sold cross borders. Errors can occur when data providers change or update systems or processes, when information is being sent to the central database and when customer data is handled poorly when entered. One of the main reasons for poor input is that workshops tend to solve information problems so that it works for the moment and for them without any concern for the central database. (Grufman)

The main sources of error are the points of data capture; processes could be automated to eliminate the human factors and thus minimising the risk for input errors.

The risk for error is highest at the sources of data entry, when integrating between systems and during input because of limitations of the front office system (Samad). According to Johansson at Ernst & Young, input errors can be avoided by having automated controls in the systems. Lindberg at Xlent lists the following as the main causes of error.
- Manual update and input of data. An update in one system can cause problems in another system, since data elements have different objectives and functions in different systems.
- Lack of knowledge of how changes in one system affect other stakeholders.

Samad explains that automated controls in the systems and standard processes for data entry are needed to reduce the risk regarding data entry errors. This can be compared to the input controls in SIS such as the correct alphabet, telephone number or email address syntax is not violated. However, as Andersson explains, there are some difficulties regarding the choice of controls. For example, there have been efforts to automate controls that restrict how workshop names can look; this has been difficult since some workshops have numbers in their name.

4.6 Fundamentals of Customer Data management

The findings from all interviews show that there are five topics that are essential when discussing Customer Data Management, Data Governance, Stewardship and Ownership, Data Quality Management, Assessment and Maintenance and Error Detection and Correction.

Data Governance

High data quality is maintained through governance. Data governance is about the processes of customer data management. Mainly, data governance is concerned with maintenance routines and distribution of responsibility. It is especially important to note that data governance is a continuous process.

Clear communication and well-defined roles and responsibilities for customer data management are the fundamental and the most challenging factors of data governance. Further, governance involves unified processes for how customer data is managed and maintained within an organisation. “If a customer walks into a workshop and needs assistance with a truck breakdown, the workshop’s focus is on giving the customer the service it needs, and fast. It is not uncommon that when the customer is not easily found, the workshop employee creates a new customer in the system, since the input of correct customer data is not a priority within the own business” (Olsson). If the customer pays with cash, the problem is solved for the moment and at a local level with disregard for how it can affect the rest of the organisation as the incorrect information is worthless for the rest of Scania. This indicates the importance of managerial communication and shared routines for how customer data is managed. (Olsson) Lovendal means that data governance routines and processes should be developed centrally and spread throughout the company.

According to Johansson and Samad at Ernst & Young, general maintenance processes and measures are developed centrally to reach unified data governance throughout the enterprise. Though, they stress the fact that it is important for the entities to be able to
chose the specifics of which data elements are measured and how, by adjusting the centrally developed processes to their business.

According to Johansson, data governance involves three steps:

1. Measure the risk of low quality on elements
2. Introduce controls to maintain high quality on high-risk elements
3. Initiate proper correction and change processes for improving data quality

Olsson stresses the fact that data governance is a continuous process, rather than a one-time fix. Kernot agrees, and states that though some one-time actions, such as mapping the business processes and applications, are needed before data governance can be implemented, data governance is about continuously maintaining high data quality.

Lovendal believes that constant monitoring and correction is needed at the source of the data.

Oatley at DeLaval explains the importance of continuous feedback regarding the governance and maintenance processes. “It is important to have a process for continuous feedback from the stakeholders.” The employees using customer data must have a communication platform to bring forward their ideas for modifying the governance processes or the central system itself, to fit their business processes and support their daily work in the most effective way. DeLaval has a change management process, which supports the users of the central CRM to give feedback to the developers of the system. (Oatley) “For example, if the French market introduces a new process, they can let us know that they will be working in a new way and need the central CRM system to support this. We will make the modifications to the system, and also see if the new process can be implemented in other markets.” (Oatley)

**Data Stewardship**

*There should be a role or an organisation responsible for customer data. There are different opinions on whether the golden copy of the customer data should be on a central or local level. Some believe that the golden copy should be placed centrally; others mean that the golden copy should be in the local systems of those that have direct contact with the customer. Where the golden copy is placed depends on where the customer data responsibility is considered to be. There are different opinions regarding if responsibility should be on a central or local level, some even suggest both. The responsibility, regardless of level, entails ensuring high data quality by supporting the stakeholders of the data.*

Roles responsible for customer data management need to be present both on a local and a central level. The central roles are responsible for the central database; they ensure that the database is functioning correctly and is easy to use. Local responsibility focuses on correct input of data. (Olsson)
The customer data in SCUD is downloaded every night from the distributors. The golden copy is considered to be the information in the local distributor systems since these are the sources of data. (Hedgren) Hedgren explains that Scania Assistance has a coordinator at each distributor in charge of SCUD maintenance that makes sure that the customer data is sent to Scania Assistance. Scania Assistance wishes to see more commitment to customer data management on the distributor level and wants to see someone responsible for managing the customer data at every dealer. Scania Assistance wishes that such a role would involve checking the local database for errors, maintain routines for quality assurance and make sure that the data is available. (Grufman) Similarly, Knoops believes that the responsibility for the customer and their data should be on a local level, meaning that every market owns its own customers and customer data. Fleet Management believes that the customer data ownership must be allocated to the local units. (Rådström) Lovendal believes that the distributors and dealers are responsible for the customer data, since they have direct contact with the customers. The routines for data governance and data error detection functions could be developed centrally. In this case, Scania Fleet Management would provide the tools for ensuring high data quality, but the data would still be owned by the distributors and dealers. (Lovendal) Currently, Scania Fleet Management has a central system solution that is suppose to support the local implementations, meaning that the golden copy of the customer data is located locally. (Rådström)

According to Pålsson, the home dealer for a customer is responsible for that customer data. The home dealer is determined by which dealer sold the customer its first product. Typically, the dealer that first sells a product to the customer remains the home dealer, even if the customer also acquires products from other dealers. In other cases, the dealers can agree on changing a customer’s home dealer. According to Pålsson there is no defined role designated for being responsible for customer data at the dealers. Customer data is typically updated when an order is made, meaning when a product or service is sold (Pålsson). Scania Mining has recently introduced a role at a few distributors, responsible for only Scania Mining’s sales. This person will also be responsible for Scania Mining’s customer data. The aim is to expand this role to other markets where Scania Mining has customers. (Winblad)

Bredenberg explains that there should be an owner of customer data for each customer, who approves all updates and changes. Unique data should be approved locally, shared data regionally or centrally, and core data changes should be approved centrally. Today, each entity is responsible for its customer information. It is imperative to convince the dealers, who see themselves as the owners of their customers, to accept that the person who is responsible for the customer data centrally needs to approve changes made in the customer data. (Bredenberg) Similarly, Lindberg suggests that some data is updated in the central database when it is updated in a local database, and some data needs to be approved before it is changed in the central database. Unlike a computer, the steward is able to discover
systematic errors and can solve it by contacting the person or unit making the mistakes and guide them into resolving the issue. (Lindberg)

Data management regarding SIS is considered successful, why it is interesting to see if data ownership is an existing issue. Andersson explains that distributors in each market have a SIS coordinator, which is responsible for ensuring that the distributors’ data in SIS is updated. Further, the SIS coordinator has continuous contact with the network of dealers and workshops in its market so that they contact the coordinator themselves when their data needs updating. Andersson, who is responsible for SIS on a central level, is responsible for ensuring that the database is functioning correctly, that the needs of all entities involved are met, and that the processes for managing SIS are in place. However, the distributors are responsible for the accuracy of the data. Further, the distributors’ data is audited every other year by a DOS coordinator to ensure that the data is correct. The focus of the audit lies on the workshops certifications for which services they offer. (Andersson)

According to Lindberg, the central custodian should be responsible for:

- Continuously monitoring data and ensuring that the data quality is acceptable
- Control and approve changes and updates made in the data that is included in the central system
- Discovering systematic errors in the system
- Knowing how data elements affect each entity
- Having continuous contact with stakeholders of the data included in the central system
- Give guidance for employees on how to follow the governance processes in place

As mentioned, Olsson discusses an issue when a new customer is created in the workshop systems when an existing customer is not found. Lindberg believes that this issue can be solved by the existence of stewards. Lindberg suggests that the service employee can register the “new” customer and the service, but also send a warning signal to the customer data steward who can investigate the matter and make sure that the information is placed where it belongs.

At DeLaval, it is believed that the ones that have direct contact with the customer own and are responsible for the data. The central employees are there to support the sales department in their management of customer data. “We only provide the tools to support their business; they are responsible for the content. It is important to be clear about that.” (Oatley)

Data Quality Management

All of the interviewees had their own definition of what data quality is, the common denominator being that the data should be good enough to not interfere with the business.
However, most of the interviewees agree that the definition and requirement of data quality should be decided by the local entity.

There is no consensus within Scania on the definition of data quality. Neither Scania Fleet Management, Scania Assistance, Scania Mining, Scania Financial Services Insurance nor Scania-Bilar Sverige has an outspoken definition of data quality. In general the different entities believe that data of good quality is correct data and is good enough to not interfere with the business. According to Olsson there are no centrally managed metrics for data quality. Of all the interviewed entities, Scania Assistance is the only one with a data quality metric.

Scania Assistance has a data quality metric called SCUD hit-rate, which measures how many percent of the customers are available in Scania Assistance’s central database. A customer post in SCUD shows if there is a customer ID for a customer and if not, then the customer information has to be input manually. Scania Assistance does not have metrics for how complete the customer post is. Scania Assistance can with a metric called SCUD Penetration see how different markets are management customer data.

According to Andersson, SIS data quality is not measured, but thinks that metrics might be interesting for the distributors and their dealer network as an indicator of how well their data management is. Olsson at Scania Networks states that data quality should be defined on a local level since quality measurements must reflect the key data elements for the business. According to Olsson, Scania uses the three C:s for data quality, meaning that data should be current, correct, and complete. Olsson believes that the definition of each term is decided on a local level, meaning every entity should define these terms to suit their business. However, he adds, the requirements need to be controlled centrally.

Samad at Ernst & Young believes that a unified and flexible conceptual model and processes for data quality throughout the enterprise must be in place. The business will need to be able to measure the dimensions that are important for their business and the definition of data quality should be decided on an entity level, since the definition should be based on the business. (Samad) Lindberg defines data quality as quality that is sufficient for the entities’ employees to carry out their daily work. “You cannot have high quality on all data elements. The entities must prioritise the elements that are essential for their business”.

Data Quality Assessment and Improvement

All interviewed entities claim to not have continuous maintenance routines. However there are routines for error correction when errors are detected. All Scania interviewees agree on that there are no central maintenance routines and that each unit handles the issue independently.

There are no centrally managed data quality maintenance routines at Scania (Olsson). In general for Scania Financial Services Insurance, Scania Assistance, Scania-Bilar Sverige and
Scania Fleet Management there are no continuous maintenance routines within the department, but there are error correction routines that come into place when error is detected in the daily business. The situation is similar at DeLaval according to Oatley. According to Andersson, SIS does not have continuous maintenance routines; however errors are corrected when detected. Scania Fleet Management and Scania Assistance have expressed wishes of potentially involving the customer in the maintenance of data while Scania-Bilar Sverige has already plans for it. Scania Mining has no structure or routines for data correction at the moment, but there is an awareness of the necessity of it for the future. Scania Fleet Management has no central maintenance routines but sees the possibility of allowing customers to themselves update their data. There have been talks about using VAT numbers to detect duplicates at Scania Fleet Management, but there are some obstacles.

At Scania Financial Services Insurance data is continuously monitored because of the frequent contact with their customers through monthly payment collection. Scania Financial Services Insurance does not have any maintenance routines involving securing the accuracy or completeness of customer data. However some of Scania Financial Services Insurance’s customer data is compared with local (accounting) systems. (Knoops) Financial services trusts that its data is correct because its customer registry is continuously reviewed due to the nature of its business.

Scania Assistance has maintenance routines involving customer data correction that are done manually and are done when errors are detected. Scania Assistance has a form of automatic error detection for when the local units upload their data to the central database, SCUD. The automatic control monitors the availability of the data. If there is a customer ID for a customer in SCUD, then that means that the information has been downloaded into SCUD, if not then the information has been entered manually because of error in the file sent to SCUD. (Hedgren) If there is no customer ID, then there is no post. Scania Assistance wishes to be able to detect more types of errors. There are pre-emptive measures taken for data quality for local databases uploading to the central database for the first time. Scania Assistance wishes to enable incremental changes in SCUD for updates, which is not possible today.

Scania Assistance sees the possibility of creating a customer portal where the customers themselves would update their customer data. Updating and taking in new data is a challenge for Scania Assistance, it is not only about being able to collect data according to (Hedgren).

Scania Fleet Management has no central maintenance routines as they feel that the maintenance responsibility lies on the creators of data, meaning the dealers.

Scania Bilar-Sverige maintenance routine involves comparing customer data with external national databases and through yearly audits of its sales related customer data. However
there are no automatic data quality controls in one of the dealers’ main local databases (Automaster), except the one that controls the organisation number.

According to Oatley there are some processes in place to prevent too much customer data being collected to begin with, since the more data is collected, the more difficult it is to keep data clean. There are no specific maintenance routines for customer data. Typically, data is corrected when errors are detected during the daily work.

According to Andersson, the SIS data is audited by the DOS coordinator which results in updates being made. The distributors have access to the SIS information and are allowed to correct error. Since the distributors are closer to the customers, it is easier for them to detect data anomalies and correct the errors.

Olsson considers the current situation with maintenance routines in Scania to be a problem since there are some customer data that might not be important for one entity, while many other entities are dependent on that data. The entities update the customer data they find essential for their business systematically.

A challenge with implementing error detection and maintenance routines is that the employees will start to see more and more errors, which can be overwhelming. Managerial communication is key, to make sure that employees understand what the definition of high data quality is. High data quality does not mean a hundred percent correct data; sufficient data quality is desired. Further, is it important for every entity to prioritise and define which data elements that are essential and should be correct. (Lindberg)

The most important issue regarding data quality is to have on-going maintenance routines, to continuously ensure high data quality (Johansson). Processes and directives should be in place for continuous maintenance (Samad).

*Error detection and controls are needed to minimise errors.*

According to Hedgren and Grufman, the SCUD system has some automatic controls. If the mandatory data elements are not sent to SCUD, the information is not uploaded to SCUD. Scania Assistance only measures the availability of customer data and not the data quality because of practical reasons (Grufman).

According to Påhlsson the dealer will launch a customer portal in the nearest future, where the customer can log on and access its information. A similar customer portal has already been launched in Scania Finland. This has been a step towards collecting a complete view of a customer. All dealers in Sweden will have access to this customer portal. It is desired to share this portal with the rest of Scania. Changes will not be made in the portal; the portal only uploads customer data from Automaster. Though, the information in the portal will mirror the customers’ internal information. In the future, Påhlsson would like it to be possible for the customer to update its information directly in the customer portal. This
would mean that the information in the customer portal must be synchronised with the Automaster. (Påhlsson) Scania Fleet Management has the idea of including the customer in the maintenance of data by asking the customer to verify their customer information through the Scania Fleet Management Customer Portal as a form of data quality control. (Lovendal; Goetzinger; Rådström).

SIS has controls for inputs in the system, but these only make sure that the Latin alphabet is used, that telephone numbers only consist of numbers, and e-mail addresses follow a certain structure. According to Andersson, there have been attempts to introduce additional controls in the system, but these have not been possible because there is no standard for other data elements. For example, some workshop names contain numbers and addresses are structured differently in various countries. (Andersson) According to Oatley, DeLaval has controls directly in the system. For example, the system recognises that a customer is Finnish; therefore the input in the telephone number field must follow the structure for Finnish telephone numbers to be saved in the central database.

Central error detection is important for when the different entities have different capacities for error detection, when information is shared and business are dependent of data managed by other entities. It is important to have continuous maintenance routines to ensure data quality.
5. Discussion

In the discussion chapter, the findings from the literature study and the empirical study are compared to reach a conclusion.

5.1 Customer Data Management

The interviews conducted at Scania show that storing data centrally has become necessary as Scania’s customers are becoming more and more multinational, which means that there is a need for sharing of customer information throughout the enterprise. One benefit of storing data centrally is that sharing of data between units and local branches is made possible. If Scania Fleet Management would share their data regarding location of vehicles with Scania Assistance, they could easily and quickly locate a vehicle that needs service. The sharing of data is highly beneficial, though it raises issues regarding definitions of data elements and data quality. For example, if Scania Financial Services Insurances share their data with Scania Assistance, they will likely have to update their data more frequently than once a month to fit Scania Assistance’s nightly updates.

The interviewees all agree on that customer data should be stored in a central database integrated with the local systems. According to Loshin (2009) a master data system includes a master data set synchronised from the original data source and made available for enterprise use; a view expressed by Scania Assistance, Networks and Olsson. A central database that is integrated with the local systems allows a unified data quality throughout the enterprise to be achieved. On a central level at Scania, the ambition is to centrally collect customer data from every market and every entity, according to Olsson. According to Olsson and Scania Networks, data needs to be divided into the three categories: core, shared and unique, to simplify the customer data management since Scania wants to store a large amount of customer data. Only core and shared data is stored and managed centrally. This view is in line with the overall concept of MDM, where only master data is stored and managed on a central level.

Scania’s central approach is supported by Samad at Ernst & Young, who also believes that customer data should be stored centrally in synchronisation with the different entities’ systems. Lindberg at Xlent agrees and explains that it is not possible for all entities to use the same system for storing customer data, since the system used must meet the requirements of a specific business. Though all interviewees agree that data should be collected centrally, there is a difference in opinions regarding what data it should contain and where the golden copy should be placed. Scania Fleet Management however, does not agree with this view of Loshin (2009). They believe that the gold copy should be placed in the local systems, since the local branches have direct contact with the customer and therefore are better suited for updating it. Scania Assistance and Scania Financial Services Insurances use a central database that better fits Loshin’s (2009) view as the departments’ ambition is for their central customer database to be a central and updated copy that collects all its information from its
local branches. Loshin (2009) claims that the integrated set of master data is the golden copy containing data that is guaranteed to be accurate and up to date.

As Lindberg explains, the data quality is improved as more local systems are involved when using the CDM approach. The more entities that use the central database, the more reliable and usable the customer data will be. Further, if many individuals use and rely on the central database, it will be given high priority in the business. Higher priority will increase the possibility of success. In addition to the central database, in interviews at Scania the possibility for allowing and encouraging customers to update their data themselves through a Customer Portal was mentioned. The main purpose of the Customer Portal will be to enable transparency towards customer by offering a portal where customers can view all of their information in one place. However, the portal can also be used within the company. The incentive for customers to update their data would be that the more information Scania has about the customer, the more customised support they can offer.

5.2 Crucial Factors and Challenges

The largest challenges when working with CDM according to the findings of the literature study and the empirical study are the following:

Organisational Rather than Technological Challenges: Both literature and empirical sources stress that while a robust technical solution is needed; the most difficult issues are organisational. Many companies make the mistake of introducing new technologies and then building maintenance processes around it. Instead, maintenance processes should be developed based on the business requirements and then suitable technologies that meet the requirements should be chosen.

Commitment: According to the literature study, the success of the introduction of CDM is highly dependent on commitment of management. Further, stakeholders must be interested and committed to the CDM efforts, since they are the ones executing the CDM activities. Managers are the drivers of the CDM effort. To create commitment amongst stakeholders, managers should communicate the benefits of successful CDM and the consequences of not prioritising the management of customer data. As Hedgren, Pålsson and Oatley mention, the more people that get involved in the CDM initiative, the more it is prioritised in the company. The managers on every level are responsible for communicating the need for CDM efforts, and create commitment.

Acceptance of Change: CDM efforts lead to changes of processes and the way employees work. Loshin (2009), and Berson and Dubov (2007) believe that this is one of the most challenging issues regarding the introduction of CDM. This issue is clear in the empirical findings as well. The consultants mention that success is dependent on how well the employees of a company can adapt to the change. The empirical study shows that the solution is managerial communication. Communication regarding why the changes made are crucial and how the benefits that it will lead to affect the employees and the company is
crucial. The importance of communication is found in the literature, where training and individual incentives are also viewed as further solutions. According to Kotter (1996), it is essential that commitment is established by identifying and discussing potential crisis or lost opportunities that can occur if change is not introduced. Further, a clear vision must be developed and widely communicated. Support is given to all employees to work towards the new initiative and the short-term gains from initiative are communicated to get even more commitment.

Communication: At Scania an issue that is clear is the lack of communication between the entities. Bredenberg explains how entities act as islands of information due to years of vertical information sharing. He explains that Scania departments have historically worked independently with customer data management. Though there are current efforts throughout the company, the efforts and ideas are often not shared or discussed between entities. Interviews at Ernst and Young revealed that this situation is not uncommon. Most entities act as independent companies when managing customer data. There are great potentials for collaboration advantages if the entities would work together and share knowledge and experience.

Reaching Consensus on Processes and Definitions: Introduction of a central customer database requires unified definitions and CDM processes. Since data is collected from independent entities, the definitions, formats and representation of the data elements will vary. Furthermore, the stakeholders, both collectors and users of customer data, have different business processes, priorities and objectives, which complicate the process of reaching consensus. The interview findings from Scania and the consultants show that reaching consensus is difficult due to that the businesses are different. The essence of the issue lies in that a customer is defined differently in various entities. The importance of shared policies and procedures for maintaining high data quality is stressed in the literature. Both literature and empirical findings show that, though the definition of customer is different, it is possible to agree on unified governance and maintenance processes of customer data by gathering expertise and opinions from different entities throughout the company. Opinions could be shared through workshops with employees from all entities affected by the customer data.

Source of Error related to Data Entry: The literature findings show that an important factor for accurate data is concerning data entry processes. Process and data standards for data entry are needed. All interviewees share the concern for input errors. Samad, similar to the literature, believes that input errors are avoided by establishing process standards for data entry. Samad and Lindberg also state that automated controls in the systems should be in place to improve error detection during data entry. This solution is used for SIS at Scania.

Data Stewardship: According to the literature, clear stewardship is essential for the success of CDM efforts. There are different opinions about on which level; central or local, of the company the stewardship should be. All interviewees view stewardship as the most crucial factor for success. Hedgren believes that there should be a central organisation responsible
for customer data. Samad and Johansson agree, and believe that CDM efforts are usually not prioritised when it is not a part of a job description. Lindberg believes that there should be clear roles for data stewardship on a central level. He means that the data is updated on a local level, since all employees should be able to update the data in their own system. Some updates could be synchronised with the central database immediately, while others need to be approved by the central customer data steward. Samad emphasises the decoupling of customer ownership and customer data stewardship. Oatley and Loshin (2009) also discuss the existence of a concern amongst the employees regarding the lack of the decoupling mentioned. Samad and Oatley, both explain that individuals are protective of their customers. Therefore, it is important to clarify that it is only the customer data stewardship that is placed centrally; the customer ownership is still local. As a contrast, some stakeholders at Scania believe that the stewardship should be at a local level, where the direct customer contact is made, which has been the case historically at Scania. Smith (2008) argues that co-ownership might be needed.

5.3 Data Governance

When viewing the crucial factors and challenges of CDM efforts, one can make clear connections between the factors and data governance principles. It is discussed that the challenges are organisational rather than technological; that an essential issue is the acceptance of change and communication between entities and executive levels. Lastly, it is discussed that it is important for clear communication; assigning of clear stewardship roles; and to reach consensus regarding processes and definitions, especially data entry processes. These issues are all connected to data governance.

Findings from the literature study are in line with the empirical findings regarding the importance of governance in CDM efforts. According to the literature, data governance includes distinguishing data governance from other types of governing organisations that exist in the enterprise, defining the objectives and scope of the efforts, deciding who is responsible for the customer data and what processes are needed to measure, analyse and improve data quality. The purpose of data governance is to maintain high data quality by developing and introducing policies, processes and activities for data quality measurement and improvement. Olsson agrees, and believes that clear communication, and well-defined roles and responsibilities for customer data management are the fundamental and the most challenging factors of data governance.

Since management of customer data is often created in isolation, formed to support functional requirements for a specific entity and without consideration for any overlaps with other businesses, the actual information architecture needs to be mapped, as discussed in the literature. This will show where customer data exists, how the data is managed and used, and how it is connected to the objectives of every entity. Scania Networks are currently working with mapping the information flow, showing where the information is stored, how it is moved from one system to another, and which entities use each system. The empirical findings show that some customer data is used by several entities, though the
data is used for different purposes. For example, Scania Financial Services Insurances register their customers insurances, which is information that is beneficial for Scania Assistance.

The literature shows that key metrics and monitoring processes need to be established to enable generation of information about the current state of the data quality. Furthermore, there needs to be standardised processes for creating, updating and deleting data. According to Johansson and Samad at Ernst & Young, this is necessary to develop a homogenous management of customer data throughout the company. The consultants believe that general process standards can be developed on a central level. Similarly, Olsson and Lovendal at Scania believe that governance involves centrally developed, unified processes for how customer data is managed and maintained within an organisation. However, the consultants add that it is important to have some flexibility for every entity to adjust the process standards to their business processes. Samad at Ernst & Young states that the entities should focus on monitoring and improving the data elements that are important for their business. Lindberg at Xlent agrees with this statement, and believes that it is not reasonable to require all entities to have excellent data quality regarding all data elements. It is most likely that the focus will lie on improving the data elements that are important for their business. Johansson at Ernst & Young believes that entities can improve their quality by measuring the risk of low quality on the data elements, introducing controls to maintain high quality on high-risk elements, and initiate proper correction and change processes for continuous data monitoring and improvement.

Lindberg at Xlent agrees with the literature when stating that there is a need for managerial communication to clarify the value and benefits of the efforts. Cervo and Allen (2011) explain that it is important to maintain data quality awareness, continue to inform stakeholders about the on-going efforts and keep the employees given the data governance support roles engaged after introduction of the data governance processes, to make the data governance a continuous part of the CDM efforts. Regular meetings should be held with the sole purpose of maintaining awareness and communicate CDM activities and the connected benefits. Interviews at Scania show that there is a strong awareness of data quality issues but a lack of organisational support.

All respondents stress that it is crucial to note that data governance and the processes it entails is not a one-time initiative. The objective is to maintain a steady state, and continuously improve data quality through on-going data quality and process evaluation, and training. Oatley at DeLaval explains the importance of a platform for continuous feedback from the users of the governance processes. The purpose is to continuously improve the processes. Feedback can lead to improvements on a local level, but also a wider level, if improvements are communicated throughout the company. Kotter (1996) believes that any change needs to be integrated into the culture to ensure its existence. The governance processes must be integrated into the way the daily work is done.
5.4 Data Stewardship

Loshin (2009) states that one of the largest challenges with data governance is the lack of follow-through. Well-defined governance policies do not lead to successful data governance unless the organisation for data governance is in place. Clear roles and responsibilities must be introduced for efficient data governance, a need that has been identified by several interviewees at Scania. Kotter (1996) states that it is important to form a group of employees that have the authorisation and competence required to lead the initiative. This is similar to the concept of introducing a central group specialised in customer data management discussed in interviewees with employees of Scania Assistance, and in the literature study. According to Karel (2007) it is important to note that governance responsibilities might be only a part of the individuals overall responsibilities.

According to Karel (2007) there should be a sponsor on an executive level to increase the potential for enterprise adoption. The sponsor should be identified early and is the driver of the CDM initiative. Furthermore, the sponsor is responsible for approving data governance policies and processes, and managing the reward frameworks. Then a team is created, composed of representatives from across the enterprise, and their responsibility is to lead and oversee data governance activities (Loshin 2009). This team, or individual, is responsible for communication between local stewards and the executive sponsor; coordination of stewards; and on-going audit of data quality and governance processes (Karel 2007). The team focuses primarily on business support and data quality management of the master data. The purpose of the stewards representing the different entities is to achieve and maintain a homogenous customer data management environment and processes. (Cervo, Allen 2011) Cervo and Allen (2011) suggest that the data stewardship model should also include local data stewards where primary creators, updaters and data users are. The local data stewards work closely together with the central stewards, and are responsible for continuously evaluating data, handle problems with data and ensure that information policies are followed. It is the steward’s task to communicate issues to the stakeholders affected. (Loshin 2009) It is important to note that these are roles and not titles. These responsibilities might be only a part of the individuals overall responsibilities (Karel 2007).

The hierarchy of the roles is shown in figure 5.1.

![Figure 5.1: Hierarchy of Data Stewardship based on the description given by Karel (2007), Loshin (2009), and Cervo and Allen (2011)]
This concept is in line with Olsson, who states that both levels of stewardship are needed. He believes that the central roles are responsible for ensuring that the central database is functioning correctly and is easy to use. Local responsibility focuses on correct input of data. At Scania Assistance, Scania Fleet Management, and Scania Financial Services Insurance, it is believed that the roles and responsibilities should lie on a local level since they have the direct contact with the customer. According to Påhlsson, it is the home dealer of a customer that is responsible for the customer data. Bredenberg explains that there should be a central role that is responsible for approving updates and changes in the central database. This role can be compared to the central steward described in the literature. In Scania’s case, unique data should be approved by a local steward at the dealer, shared data by a local steward at the distributor, and core data changes should be approved by the central steward. Lindberg agrees, and states that only a selection of data should need approval from the central steward. There are benefits of having a steward instead of computer run correction processes. According to Lindberg at Xlent, the steward would be able to discover systematic errors and contact the unit causing the errors and guide them into solving the problem, allowing the data user to not repeat the mistake.

Lessons could be learned from how SIS has data coordinators at each distributor. Scania Franchise Standards has adopted the stewardship framework by having a central steward and local coordinators or stewards at every distributor for SIS. The central steward is responsible for ensuring that the central database is functioning correctly, that the needs of all entities involved are met, and that the processes for managing SIS are in place. The local stewards are responsible for the accuracy of the data. Further, the distributors’ data is audited every other year by a coordinator to ensure that the data is correct.

As stated previously, in order to apply any responsibility structure, a distinction between customer ownership and customer data ownership must be clear for all stakeholders. This is a concern expressed by Samad at Ernst & Young, DeLaval and Scania-Bilar.

Cervo and Allen (2011) state an important issue. They explain that clear roles and responsibilities are crucial for successful CDM. However, it is important to note that an objective is to create a culture of data quality focus. Therefore, processes that enables and encourages data quality improvement should be available for all employees, and not only stewards. Further, Olsson believes that constant and clear managerial communication is needed for all employees to understand why and how CDM processes are used.

5.5 Data Quality Management

The literature findings suggests that data quality is defined differently depending on the stakeholder. The consultants at Xlent and Ernst & Young confirm this statement and add that data quality is defined by how sufficient the data is for the business to carry out its daily work. Oatley at DeLaval believes that how data is perceived depends on the role and responsibility of the evaluator. Interviews at Scania show that there are no unified definitions of data quality within the company. The literature shows that various definitions...
is not uncommon since entities usually manage customer data independently. Having different perceptions of data quality is harmless, though common conceptual definitions for data quality is needed. However, according to Olsson, data quality should be defined on a local level at every entity since quality measurements must reflect the key data elements for the business. This is important considering that the level of perceived data quality is, as revealed in the literature study, subjective.

The level of data quality that is sufficient for one application may be considered too low for other purposes. The literature study suggests that subjective assessments should reflect the needs and experiences of stakeholders. The interviewed consultants firmly believe that it is not possible to have high quality on all data elements since the definition of quality should be decided based on the business at entity level. Therefore, the data users’ experience of the organisation’s data quality becomes a key resource for identifying data quality issues. Interviews with consultants reveal that high quality data is data that is of enough high quality to minimise risks for the business.

When assessing data quality, defining the data quality dimensions is a key step in the process since the dimensions are used to evaluate the data. Batini and Scannapieca (2006) believe that most data quality dimensions should have objective measures that are not dependant on the business or experience of the user. Objective information about data can be attained through data quality indicators such as name of source and creation time. Having both objective and subjective measurements are essential for the assessments process to both comply with each department’s business needs and for finding common grounds.

As presented in the literature study, there are a number of different data quality dimensions. There are various definitions of the different data quality dimensions as authors may have various names for a dimension that essentially is the same. The time related dimensions are examples of dimensions that are defined differently by various authors. It is therefore important to identify the relevant ones for the business and to define them in a unified matter. Maintenance routines should then be developed to target the dimensions chosen.

The dimensions consistency, timeliness, uniqueness, accuracy and completeness were identified as the most common ones discussed by the reviewed authors. As Olsson mentioned, Scania aims to use the dimensions correct, current and complete, which are covered by the dimensions timeliness, accuracy and completeness. Further, from the issues expressed during interviews at Scania, it is shown that uniqueness; accuracy and completeness of data are common issues.

Pipino et al. (2002) defines the dimension of consistency as the extent to which data is presented in the same format, while Loshin (2009) sees it as to what degree data values taken from different data sets conflict with one another. Not having a standard for how data
elements should be presented may cause inconsistency in the customer data. Scania Assistance, Scania Financial Services Insurance and Scania Fleet Management have standards for the customer data structure, such as what information needs to be included in every element and how the data element should be constructed. These measures are taken to secure the consistency of the data.

The data quality dimension usually called timeliness is defined in various ways in the existing literature. However, the definitions can be summarised, and timeliness can be described as a dimension that indicates the time elapsed since the data in question was last updated. All interviewed entities at Scania update their customer data on a regular basis, with the exception of Scania Mining. Scania Assistance updates the customer data in SCUD every night, while Scania Financial Services Insurances’ update the customer data in their data warehouse at the end of each month. The different entities’ update regularity is a factor that needs to be considered for possible data sharing. For example, at Scania Financial Services Insurance a new customer appears in the central customer database at the end of the first month. If the data would be used by Scania Assistance, they would not be able to find the customer’s insurance information until the end of the month. The entities might need to make adjustments to their customer data update regularity if data is to be shared. This situation is a perfect example of how different maintenance routines present a problem as discussed by Olsson.

According to the literature, the data quality dimension uniqueness can be measured by checking for duplicates in the system. According to Kernot at Networks, duplicates do exist, something that Grufman acknowledges. However, Grufman believes that Scania Assistance’s business does not suffer from the existence of duplicates. Olsson describes a situation that could lead to duplicates, when a workshop cannot find a customer in their system and solves the problem by creating a new customer even though the customer might already exist. Duplicates might also exist since one customer can have contact with several entities at Scania, and therefore, the same customer is registered separately in the different databases.

As presented in the literature study, accuracy is the extent to which data are correct, reliable and are compatible with the values of the real-world entities it represents. Batini et al. (2009) suggests that accuracy should be measured by using a reference data set, which can be compared to what Påhlsson explains about how dealers in Sweden use the national vehicle registry database to compare their data with. Scania Financial Services Insurances does not have any on-going routines for securing the accuracy of data. However errors are often successfully detected and corrected because of the continuous customer contact which helps to improve the accuracy of the data.

Completeness is described as the degree to which data includes information that describes the real-world object it is meant to represent. Completeness is measured to some extent at Scania Assistance with SCUD hit-rate. Scania Assistance identifies completeness according to Loshin’s (2009) second suggestion, by forcing data elements to have or not to have a value.
under a certain condition. Scania Assistance also improves completeness by having mandatory data elements that must be sent to the central customer database from the distributors. Interviews at other entities of Scania revealed no methods for identifying data completeness.

Awareness of the data quality and data quality dimensions is needed in order for sharing of data not to create issues of trustworthiness for the data that is being shared and so that the data is at a level that does not interfere with the entities’ business. The entities need to be aware of that the data they create and update might also be used by another department. This is especially important for data that might not be seen as important for the entity updating it, but is crucial for other entities. Communication between the different entities is therefore of outmost importance. The central organisation must create awareness for poor data quality and the consequences of it. It is essential to have communication channels for customer data related matters so that the different entities can learn from one another or be better informed about what is being done to improve customer data management within the company.

5.6 Data Quality Assessment and Improvement

No systematic data quality assessment or improvement processes on a central level were identified during the interviews at Scania, and only a few at entity level. In most cases, data is assessed and corrected in a reactive manner, much as described by Cervo and Allen (2009) when the consequences affects the work processes of the entity. Scania Assistance has routines in place on how to correct customer data when errors are detected. The interviewed consultants at Ernst & Young emphasise the importance of having continuous maintenance routines in place. There are some routines at Scania today according to the conducted interviews: Scania Financial Services Insurances continuously compare some of their customer data with a local accounting system; Scania-Bilar Sverige uses the Swedish Transport Agency’s vehicle registry to verify vehicle ownership information; the distributors’ and dealers’ customer data is assessed through the DOS audit.

In order to address data quality issues, assessment and maintenance routines must be put in place. The assessment methodology presented in the literature study should help structure an adequate process. It is essential to include the data users in the development of the processes as suggested by the literature study, and to use subjective and objective measurements. According to the literature study, improvement of data must be preceded by a state reconstruction phase and assessment phase. The improvement steps in the literature study are specific and should be considered as potential steps.

By standardising customer data and the customer data concept, it becomes easier to compare data with each other and thus facilitate detection of errors and discrepancies. Standardised data elements are necessary in order to optimise and successfully implement other assessment and improvement techniques which of course facilitate the assessment and improvement process.
The literature stresses that the introduction of CDM efforts will lead to the detection of data quality issues. It might seem as issues are increasing but in reality, the issues have existed before, but they have been hidden. Lindberg at Xlent mentions this issue, adding that the high number of errors can be overwhelming. Issues cannot be solved if they cannot be detected. It is important for managers to communicate how high data quality is defined, since all issues that are detected might not need solving. An example is the case of duplicates in SCUD. As previously mentioned, Grufman believes that duplicates do exist but do not have a negative impact on Scania Assistances’ business.

Considering that clear responsibilities of data quality management and commitment to the cause of improving data quality on all levels is one of the main issues detected from interviews, improvement efforts should be process-driven as presented by Batini et al. (2009) as this is an organisational matter and the focus should lie in improving processes and not technology. Maintenance routines should however, as stated by the interviewed consultants, be decided on an entity level as it must be constructed so that it fits the entities’ business.
6. Conclusion

In this chapter, the conclusion and recommendations are presented. Further, a section presenting suggestions for future research is given. The recommendations are divided into the main segments of the thesis, shown in figure 6.1. The categories are based on identified issues and important factors in the existing literature, and have been supplemented with the issues and factors mentioned in the empirical study.

![Figure 6.1: The main segments of the thesis](image)

6.1 Conclusion

The empirical study shows that there is a concern regarding management of customer data. As companies grow, and customers become multinational, a well-functioning customer data management is crucial for today’s companies. Unquestionably, it is important for any company to have a clear view of its existing and potential customers. If customer data is collected correctly and maintained to keep a high data quality, it can be a great asset to the company. MDM is a concept that is used primarily for heterogeneous sources of data. It mainly involves data governance and its purpose is to collect and keep data with high quality. Using a well-adjusted MDM, or rather, CDM program, organisations can benefit all the gains that it entails.

The purpose of CDM is to obtain a unified view of a customer, which all stakeholders have access to. It is essential to note that this is a question of organisational processes, rather than technical solutions.

The empirical study revealed that there is a realisation about the importance of prioritising efforts for improvement of customer data management. Further, a concern regarding the issues caused by the lack of successful customer data management is clear in the case study. In order for a change to be introduced, top-down commitment and initiatives are needed.

Before the MDM concept is implemented, the current customer data management and current business processes of an organisation need to be mapped in detail. The CDM concepts are then adjusted to fit these processes. Stakeholders need to be involved from the beginning; the CDM processes implemented are based on the opinions of the stakeholders.
This master thesis will contribute with key insight within the area of organisational challenges that emerge when implementing the Customer Data Management concept. This thesis has supplemented Master Data Management theories with Change Management and Data Quality Assessment and Improvement theories in order to give a more overall view of common factors and issues that play part when implementing Customer Data Management. Moreover, this thesis has provided insight on key factors and concerns regarding customer data management in today’s ever evolving business and has highlighted the urgency of proper customer data management in the current globalised industry.

6.2 Recommendations

The following recommendations present how CDM can be implemented and used to collect and maintain customer data with high data quality.

Customer Data Management

In order to achieve a unified customer data management, a central database that contains updated and trustworthy customer data must be in place.

- Master data should be stored centrally to enable sharing of data and achieve better data quality
- The central database should be synchronised with the local databases, which leads to homogenous data quality throughout the company
- Customer Portal can be created to allow customers to update their data themselves

Crucial Factors and Challenges

- It is important to note that the challenges are organisational rather than technological; technology should support the CDM processes and policies, and not the opposite
- The success of the introduction of CDM is highly dependent on commitment of management on every level
- The success of CDM is highly dependent on how well the employees of a company can adapt to the change. Managerial communication regarding the importance of the change, training, and incentives are needed to simplify adaption
- There should be communication between business units regarding CDM to learn from each other and achieve a homogenous customer data management throughout the enterprise.
- Standards need to be developed for CDM policies, processes and customer data element definitions to enable integration and sharing of data
- Standardised data entry processes and automated controls in the systems are needed to minimise data entry errors
- Clear roles and responsibility for customer data is essential for the success of CDM
- A central entity responsible for CDM should be created
- There should be a clearly communicated distinction between customer ownership and customer data ownership

**Data Governance**

An organisation has to take preparatory steps before implementing data governance. A central unit for master data management and customer data management must be in place to overlook customer data management, data quality routines and the central customer database.

- Create a communication plan for whom to target, through which communication channels and with what information in order to provide clear and continuous communication from the central organisation to the business units about what the purpose of the data governance is and the consequences of poor customer data quality to create awareness and commitment for the efforts. The more individuals and business units that use the central database, the better the data quality will be, and the higher it will be prioritised in the daily work
- Introduce regular meetings with the sole purpose of maintaining awareness and communicate CDM activities and the connected benefits
- Map the customer information to see where and what information overlap exists and what customer data is being used by which departments and business units.
- General maintenance processes and measurements are developed on a central level but should be flexible enough for business units to adjust them to their business processes
- Governance efforts are not a one-time initiative and require continuous efforts as they need to become a part of business processes and thereby, everyday work

**Stewardship**

There must be clear customer data management roles with clear responsibilities such as customer data stewards on different levels of the organisation.

- Stewardship needs to be present on different levels an organisation
- A steward on an executive level should be in place, and should be responsible for providing strategic direction, reviewing policies and assigning stewards to define information policies and rules based on business policies, approving data governance policies and processes, and managing the reward frameworks. This role is preferred but not necessary; these responsibilities could be given to the central stewards
- Create a group consisting of stewards from different business units, on a central level with the primary focus on monitoring, maintaining and approving updates of the master data stored in the central database. Furthermore, the team of stewards should be responsible for giving support regarding customer data management the lower level stewards and other stakeholders. This could consist of a group or one individual depending on the needs of the organisation
- Create data stewardship roles at different levels of the organisation with clear customer data responsibilities
- It is important to note that these roles do not have to be full-time. The needs of the organisation determines the required number of hours and employees needed for the customer data management roles

**Data Quality Management**

Data Quality Management is only successful if many stakeholders are on board on the matter and there is a willingness to cooperate across the organisation.

- The possibilities for sharing customer data should be communicated between an organisation’s entities
- The consequences of poor data quality management should be communicated within the organisation so that the different departments understand the needs of other stakeholders
- Map what data is being used by which departments and determine which dimensions are critical to the different departments of the organisation.
- Standard definitions for data quality dimensions should be defined, and valid for all parts of the organisation
- Definition of data quality on a conceptual level should be defined at a central level while the definition of what is sufficient quality should be decided on a business unit level.

**Data Quality Assessment and Maintenance**

Awareness about poor customer data quality and the consequences of it amongst stakeholders is important in order for data quality issues to be resolved and improvement measures taken.

- An assessment process with adequate data quality improvement activities should be developed with room for adjustments for the different business units of an organisation
- Customer data quality should satisfy the business units’ data requirements
- Maintenance and improvement routines suited for the business units’ data requirements should be in place at each business unit
- Data quality should be assured by the business units to meet the business units’ data quality requirements
- Data Quality goals should be set up centrally, created from the assessment process with both objective measurements and subjective measurements

**6.3 Limitations and Future Research**

For future research, studies of other cases involving a central customer database conducted as a continuation of this study would give more general results. Studying organisations that
have implemented MDM to a higher degree could give a better indication on critical issues that can appear during the process of implementation.

Considering that this study has established that a change in how customer data is managed is necessary, it would be beneficial further research on change management and how it can be used to improve CDM efforts.

It would be beneficial to complement this thesis by doing additional research in the area of data governance and the assignment of responsibilities for data quality assurance as the thesis merely gives an overview of important factors regarding CDM. Future research could provide models for governance and stewardship on a central and business unit level for organisations using MDM. This thesis only provides a general data quality and improvement processes which would need to be adapted to an organisation’s needs and business. In order to create models for data quality assessment and improvement for a specific organisation, it is advised to use this thesis as the foundation for such research and in addition take under consideration the specific organisation’s policies and processes.
References

Publications


Electronic

Britannica:
www.britannica.com (A) URL:
http://www.britannica.com/EBchecked/topic/155435/deduction
Collected: 2012-02-20

www.britannica.com (B) URL:
http://www.britannica.com/EBchecked/topic/286724/induction
Collected: 2012-02-20

Ernst & Young:
Collected: 2012-03-19

Xlent:


Scania Bilar Intranet:

Collected: 2012-04-20

Scania Group Homepage:


Scania Inline:

http://inline.scania.com (A), URL:

http://inline.scania.com (B), URL:

http://inline.scania.com (C), URL:

http://inline.scania.com (D), URL:
Collected: 2012-04-05

Search Data Management:

http://searchdatamanagement.techtarget.com,
URL: http://searchdatamanagement.techtarget.com/expert/David-Loshin
Collected: 2012-04-15
# Appendix

## Appendix A: Wordlist

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>A computer software designed to help the user to perform specific tasks.</td>
</tr>
<tr>
<td>Dealer</td>
<td>Facility within Scania’s authorised distribution network of products and services, in charge of sales and distribution of new Scania products and after-sales services. The facility includes a sales office where the sales force is stationed and where the customer can meet a salesman on a regular basis. 1. When a facility sells Scania products in its own company name, it must comply with all requirements in the Scania Dealer Operating Standard. 2. When a facility sells Scania products in another company name, it must comply with all requirements in the Scania Dealer Operating Standard with exception of the specific requirements that are relevant and limited to sales and distribution of new Scania products.</td>
</tr>
<tr>
<td>CRM system</td>
<td>Customer Relations Management System</td>
</tr>
<tr>
<td>Department</td>
<td>Central Scania organisation such as Scania Assistance, Scania Fleet Management and Scania Financial Services</td>
</tr>
<tr>
<td>Distributor</td>
<td>Member of Scania’s distribution network of products and services, in charge of wholesale and the distribution network for one or more of Scania’s products and services in one or more countries</td>
</tr>
<tr>
<td>DMS</td>
<td>Dealer Management System. IS/IT-system used by a dealer/workshop to administrate the distribution of products and after-sales services</td>
</tr>
<tr>
<td>DOS</td>
<td>Dealer Operating Standard, an internal standard that specifies the minimum requirements that a member of Scania’s authorised distribution and service network must meet</td>
</tr>
<tr>
<td>Entity</td>
<td>Overall term that includes dealers, distributors, departments and business units</td>
</tr>
<tr>
<td>ERP</td>
<td>Enterprise Resource Planer</td>
</tr>
<tr>
<td>FMS</td>
<td>Fleet Management System</td>
</tr>
<tr>
<td>Golden Copy</td>
<td>Central copy of most updated version of master data</td>
</tr>
<tr>
<td>Information Architect</td>
<td>Person who's work competence is in the area of information architecture which is the science of organising and labelling websites, intranets and software to support usability</td>
</tr>
<tr>
<td>SCUD</td>
<td>Scania Customer Database, Scania Assistance's central customer database</td>
</tr>
<tr>
<td>SIS</td>
<td>Scania International Service</td>
</tr>
<tr>
<td>VAT Number</td>
<td>Value Added Tax Identification Number, an identifier used in the European Union</td>
</tr>
<tr>
<td>Workshop</td>
<td>Facility within Scania’s distribution network of product and services, exclusively in charge of after-sales services</td>
</tr>
</tbody>
</table>