UNIVERSITY
OF TURKU

Master data management in industry

Master of Science in Technology Thesis
Software Engineering
Degree Programme in Information and Communication Technology (MSc. Tech.)
Department of Computing, Faculty of Technology
University of Turku

Author:
Arttu Nurminen

Supervisors:
Antero Järvi
Anne-Maarit Majanoja

24.4.2022

The purpose of master data management (MDM), is to create and maintain company's business core data. In this procedure, information technology processes and business processes are integrated into each other. Master data is used for example in product information, employee information, customer information and supplier information. Master data is a base for using transactional data due to for example sending of invoices need customer information which belongs to master data.

In this research, MDM was researched in the branch of industry. The theoretical part combined a base from publications and literature. In addition, the improvement of the MDM in the case company was considered.

As the main findings of the theoretical part were listed: better MDM mirrors to decreasing of redundant costs, better decision making with more reliable data, deviant data mastery policies in MDM compared to normal data mastery policies, MDM's company-wide scope and the important role of MDM in possible changes of business needs.

Implementing MDM in industry includes three steps which are found to be the most important ones. These steps are: master data model planning, implementation style selection and master data change request planning.

The case company had challenges with MDM which occurred mainly from using Excel files which made the management less reliable and these files congested mail servers. In addition, the process included unnecessary manual interphases.

The new MDM solution for the case company made the MDM more reliable. The improvement contained a centralized database and a web application. This removed the need for using Excel files in MDM. The solution also enabled the master data to be transferred into downstream applications with less manual interphases.

Key words: master data, master data management, data architecture, data quality

Diplomityö
Tietotekniikan laitos, Teknillinen tiedekunta
Turun yliopisto

Master datan hallinnan tarkoituksena on luoda ja ylläpitää yrityksen liiketoimintojen ydindataa. Menetelmässä integroidaan tietotekniset ja liiketoiminnalliset prosessit toisiinsa. Master dataa käytetään esimerkiksi tuotetiedoissa, työntekijätiedoissa, asiakastiedoissa ja toimittajatiedoissa. Master data on perusta transaktionaalisen datan käytölle sillä esimerkiksi laskujen lähettämiseen tarvitaan asiakastietoja jotka ovat master dataa.

Tässä tutkimuksessa tutkittiin master datan hallintaa teollisuuden toimialalla. Teoreettinen osuus kokosi perustan julkaisuista ja kirjallisuudesta. Lisäksi tarkasteltiin master datan hallinnan parantamista kohdeyrityksessä.

Tutkimuksen teoreettisen osuuden päälöydöiksi lueteltiin: paremman master datan hallinnan heijastuminen tarpeettomien kustannusten vähenemiseen, luotettavamman datan perusteella tehtävien päätösten paraneminen, normaalista datanhallinnasta poikkeavat politiikat master datan hallinnassa, master datan hallinnan yrityksen laajuinen ulottuvuus ja master datan hallinnan tärkeä rooli mahdollisissa liiketoiminnan tarpeiden muutoksissa.

Master datan hallinnan toteutus teollisuudessa sisältää kolme vaihetta joiden on todettu olevan tärkeimpiä. Nämä vaiheet ovat master data mallin suunnittelu, toteutustyylin valinta ja master datan muutospyynnön suunnittelu.

Kohdeyrityksellä oli haasteita master datan hallinnan kanssa mikä johtui pääasiassa Excel-tiedostojen käyttämisestä. Näiden tiedostojen käyttäminen teki hallinnasta vähemmän luotettavaa ja ruuhkautti sähköpostipalvelimia. Lisäksi, prosessi sisälsi tarpeettomia manuaalisia välivaiheita.

Uusi master datan hallintaratkaisu teki kohdeyrityksen master datan hallinnasta luotettavampaa. Parannus sisälsi keskitetyn tietokannan ja web-sovelluksen. Tämä poisti tarpeen käyttää Excel-tiedostoja master datan hallinnassa. Ratkaisu myös mahdollisti master datan siirtämisen loppupään sovelluksiin vähemmillä manuaalisilla välivaiheilla.

Avainsanat: master data, master datan hallinta, data-arkkitehtuuri, datan laatu

Contents

FIGURES

TABLES

1   Introduction

Master Data Management (MDM) means a way with which an organization manages data which is used in multiple locations. In other words, the same data can exist in many locations but the data must remain coherent in all locations although the data changes. Proper MDM process is good to be emphasized at the latest when the organization grows. Smart using of master data in company mergers and acquisitions creates synergy advantages. If an organization is a company which for example sells products, so the product prices must be the same in every system within the company. Many companies have not thought the master data management before challenges have started to occur. It is more difficult to create operational master data management subsequently, in other words it takes more human work to modify processes and systems to work proficiently comparing to situation where functional master data management has been taken into account as soon as possible.

Managing master data is not only an information technology challenge but it is also needed to achieve business goals. This requires continuous governance to be able to use master data as efficiently as possible in decision making. Poor master data management in turn leads to direct and indirect costs. In addition, costs in form of fines can occur if company does not manage data as it is required in existing law so it is also a responsibility to manage some defined master data properly.

1.1   The objective and research questions

This master's thesis was written as a part of the master's degree of information technology in the University of Turku. The thesis combines theoretical research to be used in practice.

The purpose was to research MDM in theoretical context and then implement a new master data management solution for the case company. The goal of the new MDM solution was to resolve challenges of current solution with new process and system. The solution was implemented for Teleste PLC (referred in this thesis as Teleste). Teleste is an international technology corporation which was founded in 1954 and it is listed on Nasdaq Helsinki. The corporation offers video and broadband, security and information technology and related services. [12]

This thesis considers the following research questions:

RQ1. What must be taken into account when implementing master data management in industry?

RQ2. How the current master data management process in the case company can be improved to fulfill predefined requirements?

RQ3. What kind of technical solution the new master data management process will need in the case company?

1.2   This thesis

This thesis has been divided into five chapters (figure 1.). These chapters form three main parts. The first part considers the theoretical background and it includes two chapters which are Introduction to MDM and MDM industrial implementation and concepts. The second part considers the MDM implementation for the case company and it includes two chapters which are MDM challenges in the case company and New MDM solution for the case company. The two mentioned parts are then summarized in the Conclusions chapter.

Introduction to MDM → MDM industrial implementation and concepts → MDM challenges in the case company → New MDM solution for the case company → Conclusions

Figure 1. Thesis structure.

1.2.1   Theoretical part

The theoretical part considers the MDM from two point of views. This part includes two chapters which are Introduction to MDM and MDM industrial implementation and concepts.

The literature review considers topics which have been emphasized in research in the context of master data management. The review has been divided into three sections. The review starts with opening what the MDM concept means. The next section considers why proper MDM is important for organizations. The third section considers how organizations can implement own MDM and what is good to be taken into account.

The MDM industrial implementation and concepts chapter considers central concepts which belong to MDM implementation in industry. These topics are model creation, different implementation styles and change request creation.

1.2.2   Case part

The case part considers the hands-on MDM implementation for the case company. This part includes two chapters which are MDM challenges in the case company and New MDM solution for the case company.

The MDM challenges in the case company chapter describes what kinds of challenges were faced with MDM and which were the reasons which created the challenges. These challenges created a need to start planning a new MDM solution.

The new MDM solution for the case company chapter considers the requirement specification, implementation, deployment and validating the solution. The requirement specification was written based on the faced challenges which wanted to be resolved. The implementation section considers what kind of solution was implemented. Validating the solution section considers the validation of the new solution mirroring to the requirement specification.

2   Introduction to MDM

This chapter considers topics which have been emphasized in researches in the context of master data management. The chapter has been divided into three sections.

The researches have been gathered from publication archives of IEEE Xplore, University of Turku, Tampere University of Technology and Lappeenranta University of Technology. The used search clauses were: MDM, master data management, master data management MDM. Altogether 12 researches were selected to be considered more closely.

2.1   Master data management

This section has been divided into three parts. The first part considers what kind of different features are related on master data management. The second part considers master data in companies' operations. The last part opens the relation between master data and transactional data.

2.1.1   Features of master data management

Companies have different kind of data and usually certain data can be found more important than other data. The data which is the base for company's key business processes and systems are considered as master data. This means that for example if a specific data attribute exists in eight of ten business processes, this data is likely master data. Same master data exists usually in many systems due to for instance data architecture, network topology, system design, geographic constraints and transaction speed. Using distributed data means that there must be a way with which the data must be kept updated in all systems. [5, 15, 16]

Purpose of the master data management concept is to resolve quality challenges of data in its whole life cycle. In other words, the purpose is to ensure long-term master data consistency although the same data exists in many systems. Companies often divide master data into entities which can be for example products, contracts, accounts, employees, customers and suppliers. Thus, master data describes company's core entities. Master data management focuses solutions to problems which are due to incoherent processes, independent systems, complicated architectures and data fragmentation. Master data management can be done with dividing the management into two parts. The first part is business processes and the second part is standardization and integration of information systems. [1, 7]

It is worth to spend resources into master data management because a little quantity of erroneous master data can result in notable costs for the company. The challenge is often that it can be challenging to govern the entity because master data is a company-wide resource. This scope is the thing how master data management differs from the traditional data management practices. [1]

Companies facing changes like mergers or acquisitions usually result in data mastery projects which contain for instance merging same kind of systems to each other. These often handle master data and it is a good place to allocate resources to implement more coherent master data management. [7]

The technical master data management system contains combined, cleaned, standardized and rationalized data. This data is also called a golden record. Attributes of the golden record are for example accuracy, relevancy, timeliness, completeness and accessibility. This data can be for instance broadcasted to other systems whereupon the data will remain updated in all systems. However, this is not always an easy task because sometimes different systems do not have all required fields whereupon there can show up empty field values when building up the golden records. Empty fields will cause challenges for example in corporate reporting. On the other hand, if the data is started to be supplemented manually, the reports can become distorted if for instance part of the added data is newer than other data, in other words the added data is not gathered from the same time period due to the lack of timestamps. [1, 6, 7, 14]

Master data management does not mean only that the company stores a lot of data into database and governs this data. Proper master data management means that a company understands which data has the biggest effect on business. In other words, it means that the management is considered more about quality than quantity. Correct master data is a precondition for company's performance. Companies have different kinds of functions and the master data management is one of the important ones because it manages the company's common core data which is used in business processes and also verifies that this data is available. The continuous availability is one of the reasons why implementing master data management is a continuous cross-functional activity instead of a temporary project. Its cross-functional role can be compared to enterprise resource planning which manages cross-functional business processes. In order that the mentioned function will work, it needs

employees which have right roles and right responsibilities and finally it will need technology as an enabler. [1]

Master data management is not only an information technology project but a complex operation which has many dimensions because it integrates business and information technology operations where the same master data is used in different systems by different company groups. Thus, master data management differentiates from normal data mastery policies due to its company-wide scope. Master data management can be also regarded as an enabler for company-wide information governance program. [1, 6]

Master data needs continuous and proper data governance. This means that the governance must monitor the quality periodically that the master data meets the needs of the business. The goals are preferable decision making, compliance verification, business performance growing and business integration contribution. The goals will be defined as strategic, tactical and operational policies which the company have to obey. The master data governance includes resource allocation where different responsibilities and tasks will be specified for concerning organizational roles. [1]

Practically, resolving master data quality challenges can be done with recognizing incorrect data which can be already in the systems and restricting data which is coming from user for example via input fields. Incorrect data in system causes often different kinds of symptoms which depend on how the data is incorrect. Regardless of incorrect data, a system can continue functioning normally. This kind of situation occurs for instance when some value has a misspelling, some query returns wrong values or some query returns only a part of the desired data set. Incorrect information can cause visible error messages but incorrect data can also cause the system to stop functioning. [4]

A sufficient quality level of master data depends also on the target users. Public and private organizations have different requirements by the country's government. Public organizations might be obligated to publish some data as open data which creates stronger requirements for data usability and that the data must be up to date and correct. In situation where data is available through application programming interface, it is important that the data will be in correct form in order that it can be handled automatically in subscriber's systems. It is important to let the subscribers know early if the structure of delivered data will be changed. [1]

Getting the quality of master data into sufficient level can vary depending on how the data is hosted. For example, in cases where master data is managed fully or partially with cloud computing, in other words hosting applications and data outside of company's own data centers, produces more challenges. This is a consequence of having less control on integrations for instance. [7]

2.1.2 Master data in companies' operations

Due to the fact that master data is used by many business operations, the creation of data attributes must be thought also from many viewpoints and use scenarios. Different business operations appreciate master data attributes which are more essential to them whereupon different attributes might be more important and less important at the same time. In order that the master data would be equally beneficial for every business operation, the aim of master data management is to create a comprehensive 360-degree view of the master data. This kind of view also reveals different relationships in respect of stakeholders. A specific customer might be a customer for many business units of the enterprise. In addition, a specific customer might be also a supplier. The mentioned relationships are good to be able to reveal quickly for perceiving a big picture when for example a management team makes strategic decisions. Decision making is easier when for instance the lifetime value of the customer can be calculated and based on this information the marketing can be better focused on the most important customers. [5]

A big picture can be more difficult to form without proper master data management because different systems tend to handle and save records a little bit differently. Especially in company mergers and acquisitions where companies have same customers or same products might lead to situation where company group's management team does not know easily for example how much a specific product has been sold or which are the most important customers. This kind of connection challenge is often handled with matching attributes in master data management. Gathering all pieces of necessary data is the key for creation of a comprehensive base which is needed when transforming data into information and information into useful knowledge. [5]

One of the concrete processes where master data management can be demonstrated is the order to cash process of a product. To be able to offer a product to customer and deliver the product successfully, the company must handle attributes which are for instance basic product properties, stock availability, estimated delivery times, prices for different package sizes,

invoicing terms and delivery address. The mentioned attributes have been usually spread in many systems whereupon data management has to work properly to be able to keep the customer satisfied. [7]

2.1.3 Master data and transactional data

Master data is static and data modifications happen more sparsely compared to for instance transactional data. Master data is the base for transactional data because for example invoices need customer name and address which are master data, in order that the invoices can be sent. [5]

Master data and transactional data have some properties which differ significantly compared to each other. These properties are related to how the data is spread, how the data is upheld, how notable the data is, how often the data is changed, what is the origin of the data, how many data items exist in the long run and how independent the data is (table 1.). [5]

Table 1. Master data and transactional data differences (table modified based on [5]).

	Master data	Transactional data
Spreading	Broadly spread	Mostly local and tied to designated system
Upholding	Upheld in multiple systems	Upheld locally in designated system
Notability	Extremely notable	Not very notable
Rate of change	Not often changed	Continuous change
Origin	Stems from genuine world	Stems from inside organization
Count in the long run	Count of items pretty invariable	Count of items varies continuously
Independence	Completely independent	Existence requires master data

2.2 Importance of proper master data managing

This section has been divided into five parts. The first part considers background for increased importance for master data management. The second part considers master data in business. The third part considers how master data relates on companies' decision making. The fourth

part categorizes common master data challenges. The last part opens how poor master data quality effects in different company layers.

2.2.1 Background for increased importance of master data management

Decades earlier when computers were becoming available to companies, the master data was kept in a couple of computers and usually the same data was kept only in one computer. The data maintenance was much easier at that time. In the course of time, a proper master data management has become more important due to the amounts of systems and data has been increased. Using the same data in many systems makes the maintenance challenging and expensive. It is common that the level of data management has not been kept up with the increase in the amount of data. [1]

Company functions change and improve to be able to keep the competitiveness at short-term and to survive at long-term. Master data management must be taken into account in changes more and more because changes are often linked to tasks, employees, organizational structures and leadership processes. The mentioned linked things include a lot of master data so it eases the changes when master data has been managed correctly. [1]

The way that companies handle sensitive master data is a competitive advantage especially in private sector. Sensitive data management procedures can be told to clients openly for example in organization's web sites or newsletters. Clients emphasize certain features when considering purchases and proper sensitive data management is one thing which weight increases all the time. In addition, development of master data management practices can make a powerful competitive advantage against competing companies. Due to increasing awareness of advantages of master data management, data-driven business models are becoming more popular. These models can be easily adapted to the market needs. [1]

2.2.2 Master data in business

Master data forms a basis for business information and due to this, it is essential for business. Without proper master data management, the data might become distorted which will show up as duplicates, missing attribute values and data value conflicts. Distorted data and incoherences result in qualitative and financial losses. For instance, incorrect product data, customer data or payment terms data cause problems in business events because for example

sales orders consist of this information. High quality data is also an end-product which means that the data must be seen also from business aspect, not only from technical aspect. [1, 2]

Master data management improves core strategic, tactical and operational processes. It also takes care of data quality of other systems such as customer relationship management or enterprise resource planning. Proper master data management also improves company's expertise image. [2]

Value of well managed master data can be easily measured also with other entities, for example with product data management. Poor product data management might lead for instance to situation where product count in stock does not hold true. Poor customer data management might result in situation where customers will not receive their ordered products on time due to incorrect or insufficient shipping addresses. [7]

It is common that only a little part of organization's data can be defined as master data. Due to the importance of master data quality in business processes, the resources of data quality enhancement endeavors are worth to focus on master data. [5]

Generally, poor data quality will lead to direct and indirect costs. Direct costs mean for instance using resources to fix incorrect data. Indirect costs mean for example improper decision making or reputation harm. [2]

Managing data properly is also a responsibility for companies. For instance, data privacy laws enjoin how data must be handled and stored. Poor data quality can lead also to situation where company's accounting does not hold true which can lead to serious consequences. [2]

2.2.3 Master data in companies' decision making

Master data combined with transactional data are used to tell executive managers how well the company is functioning from different perspectives. This is often performed with business intelligence -tools which are used for instance to enrich the data to information and form informative reports which can be used in decision making. This is an example where analyzing of unnoticeable wrong data can lead to wrong decisions which have concrete financial consequences. There exist also worse scenarios if for example automated systems are used in decision making whereupon it might take a long time until some human notice that something is not correct, not to mention financial loss which can be very huge within a small time. Detected wrong data without quick correction opportunities leads to situation where the

data must be retrieved manually which takes significantly more time before decisions can be made. [6, 9]

Although the data is used as a basis for many decisions in organizations, there is rarely used enough resources to maintain the quality of data. Sometimes there is no designated person who should be responsible for taking care of data quality. This might lead to situation where people think that the data quality will stay correct because there is someone whose responsibility is to take care of the data quality. This kind of false belief is often due to opaque processes. [2]

With proper master data management, there is no need to spend time to find which data is the correct one if different versions of same data appeared in different places. Easy finding of the single version of truth makes the decision making easier because for example many business decisions have been made based on reporting which has been consisted of a combination of master data and transactional data. In other words, trustworthiness of financial data is vitally important for companies. [2]

2.2.4 Categorization of common master data challenges

Challenges with master data can be categorized with different types. There can be seen five main types [4], which occur most:

- Wrong data

- Broken data

- Unreachable data

- Systems' lifespans

- Insufficient system design

The first challenge is that a system includes wrong data which for instance can lead to a situation where a delivery to customer is sent to wrong address due to outdated information. The second challenge is that a data can be broken which means for example that input fields do not sanitize data which user gives, whereupon data processing might have to be done manually due to detected errors. The third challenge is that a data can't be reached which kind of situation can occur when for instance an intermediate system is unavailable which leads

often to situation where part of data is missing. The fourth challenge is that systems have lifespans, new system will be integrated to current systems, migrations will be done between systems and this all makes the maintenance of master data more difficult. The fifth challenge is that a system design has not taken into account a situation where for instance a system allows to store only one address for one customer company but there has come a need to store more addresses for example a bill-to address and an order ship-to address which might lead to situation where same company has been stored many times in database with different addresses. [4]

2.2.5 Effects of poor master data quality in different company layers

The effects of poor master data quality mirrors into multiple layers in a company. Three main layers of these are strategic layer, tactical layer and operational layer. Each layer can be divided into consequences caused by poor data quality (table 2). [5]

Table 2. Poor master data quality effects in strategic, tactical and operational layers (table modified based on [5]).

Layer	Effect
Strategic layer	Challenges in specifying and implementing strategy Worsening of company's cohesion culture
Tactical layer	Weak decision making Challenges in data warehouse expansions Rising lack of confidence inside company
Operational layer	Client discontentment Wasted potential sales Rising expenses Decreased working contentment

On strategic layer, poor data quality can have the utmost-reaching impacts in all three mentioned layers. It can hinder strategy phases because analyzing and adjusting strategies with inaccurate data does not lead to best results and the strategy modifications are often more difficult and more expensive afterwards. Challenges in implementing strategies can result in bad debates among employees worsening the cohesion culture of the company. [5]

On tactical layer, poor data quality can have also far-reaching impacts but less than on strategic layer. This is due to the fact that when the decisions will be made based on the data then the quality of decisions cannot be better than the quality of data. Although the poor quality of data is known it reduces the speed of decision making. The data which is used in decision making is often gathered from data warehouses where can occur challenges when the data warehouses are wanted to be expanded in order that to make better decisions which are based on larger amount of data. Noticing the poor quality of data can lead to rising lack of confidence to other business units' expertise. This kind of lack of confidence can cause a situation where different business units create and maintain their own data silos which wrecks the idea of mastering data as an entity. [5]

On operational layer, poor data quality might result in client discontentment due to for example erroneous shipment or erroneous billing. This means that clients must spend additional time to resolve the challenges which naturally increases discontentment. Insufficient data wastes potential sales which would have been noticed with better data. These potential sales can be a result from noticing new trends or noticing more closely the things that clients really need. It is also possible that company does not know all the clients because systems might return different client lists whereupon business management is much more difficult. Examining the reasons of poor data quality and fixing the situation requires resources which rises expenses. Inaccurate data does not reveal certain expenses which a company would decrease for instance with putting specific resource purchasing out to tender. Client discontentment mirrors to decreased working contentment because employees must handle poor data quality related issues with discontented clients. [5, 6]

2.3 Implementing master data management

This section has been divided into four parts. The first part considers starting to implement master data management. The second part considers master data management in different company sizes. The third part considers the importance of documenting master data management. The last part considers the evaluation of master data.

2.3.1 Starting to implement master data management

Organization's master data management can be utilized with practices which have been found beneficial but there is not one universal recipe which would fit to all organizations. All organizations have to develop their own master data management styles to be able to create,

maintain and utilize the master data best. Although proper styles have been created, the organization's management team has to make sure that all defined stakeholders know the rules and obey them. [5, 7]

It is not worth to rush into managing master data with new approaches before existing data challenges have been understood completely. The challenges of implementing master data management are often related more to governance than technical issues. This applies also with other systems such as enterprise resource planning and customer relationship management. [7]

Data has been often seen as a thing which quality can be improved only with a new system. Improvement of data quality requires also that processes must be improved because the original reason for improving master data management is a business problem which cannot be improved only with information technology. In addition, successful implementation of master data management needs also co-operation with corporate finance department. [2, 7]

Preparations for possible master data management project includes finding out the value which it will create, how much the project will cost but also the possible costs if the project will not be started. The value can be for example business case related problems which the new processes and the new systems are intended to resolve. Direct and indirect costs which originate from poor data maintenance are a good metrics for estimating consequences if the project will not be started. This will also express how urgent it will be to start the project. [6, 7]

After preparations, the benefits of the changes are important to explain understandably to governance because merely the changing of process needs resources already not to mention investments to new technological tools which are often needed in addition to process improvement. Possible harms of not implementing the master data management properly are good to be emphasized in order that possible risks of not implementing the project are understood well. [6]

Master data management project is useful to begin from areas most crucially influenced by master data issues. Global master data management solutions must take into account flexibility of local rules more strongly because for example some subsidiaries might need different kinds of pricing rules. In addition, possible company mergers and acquisitions in the future are good to be take into account always. [7]

Some organizations might create or update same objects in different systems separately whereupon misspellings might lead to confusions. This also takes more time to keep data updated. To resolve this challenge, it is important that the master data management will be planned with a way where the master data must be created and updated only in one system and then the data is subscribed by other systems. [2]

Starting to manage master data should be done with moving forward step by step in the roadmap than at one time. With this way it does not bring down the whole big project if some planned part does not work as expected. In other words, every step should be tested at a time which also saves time because the testing can be done earlier than testing the whole final result in the end. [3]

Master data specification is among the first steps when starting to manage master data. There needs to be examined which objects in business processes can be identified as master data. The source of every master data object must be found out. It is very likely that there exist same data objects with different representation in different systems. For example, last names and first names can be in different fields in one system but in other system they can be a comma separated combined value in one field. Each master data object must have a primary business owner who takes care of the whole data life cycle. It is good to emphasize the business word because the data owner should rather be from business than from IT. [2]

Implementing master data management would be worth to be started from the source of the data. Usually if the data will be fixed in the subscribing systems, the data must be fixed also in other systems which subscribe the same data. It is important to check periodically that the broadcasted data has been copied correctly into the subscribing systems. Automatic error messages are worth to be built into the copying process. However, this is not always an easy job because checking data automatically needs familiarization with the topic. Challenges related on this task can occur for instance with finding duplicates automatically because there might be some little differences between records whereupon these records are not recognized if searched by exact copies. To resolve this challenge, approximate matching process needs to be implemented and for example a fuzzy search could be an efficient ancillary in this case. [4, 5, 14]

2.3.2 Master data management in different company sizes

The easiness of managing master data depends a lot on the size of the company (table 3.).
Often it is easier to manage master data in small companies. Especially when a company is in
its early stage, managing master data has not been usually taken into account as much as it
should be before problems arise which usually happens for example when the company
grows. Young companies usually pay attention more to solutions like enterprise resource
planning or customer relation management. Medium size companies are starting to awake to
challenges originated from the growth of data amount. To solve the data challenges, medium
size companies determine data stewards. Large and conglomerate companies have often more
data and more systems which increase the complexity. Conglomerate companies with many
businesses have to handle also a lot of master data groups at the same time. In addition, large
and conglomerate companies which operate globally have to deal with both internal and
external factors which vary depending on the residing market, meaning that there are often
different requirements for data quality. It is important to define the balance between how
much data must be handled with a globally standardized way and how much there exists local
flexibility. It is also important to plan the master data management to be flexible if changes
must be done in the future. [6]

Table 3. Characteristics and central challenges in different company sizes (table modified based on
[6]).

Company size	Characteristics	Central challenge
Small	Little quantity of master data Integration not priority due to simple implementation	Scaling plans for future
Medium	Data stewards determined due to integrations are becoming challenging	Creating efficient controls and standardization
Large	Large quantity of master data and different kinds of system integrations	Achieving unanimity between essential stakeholders Administering all created integrations
Conglomerate	A lot of master data groups made by different kinds of businesses	Specifying the general level of master data

Fixing wrong master data becomes more time-consuming without proper management when the same data is used more in different systems which can be the situation for example when a company grows. Company mergers and acquisitions make the master data management more difficult due to that different companies often use different data architectures. Especially the level of difficulty increases if the companies are planning to combine their actual functions on mergers or acquisitions. To resolve this situation, it would be better to improve the companies' data architectures to uniform solution. This also yields synergy advantages. [3, 6]

Companies often have not time or interest to examine functions which seem to work without insurmountable problems. This is emphasized especially with functions which do not have designated owners. It is common that while implementing any kind of data improvement project, there might be revealed functions which have worked wrong all the time but everybody has trusted that it has worked correctly and the consequence has been sufficiently little that nobody has not suspected any malfunction. [6]

2.3.3 Documenting master data management

A detailed and up-to-date documentation is important in master data management. The documentation is especially important when the master data administrators change. Sometimes a data model change updates to documentation might be forgotten and this does not have an immediate impact on master data maintenance but when time has lapsed and some new data model changes must be done, then there might come big problems. Due to this every little change in data model must be documented although the change would be small. It is also important to document the data model changes within as little time as possible because otherwise the probability to forget to document the changes increases when time has lapsed. Sometimes master data can be removed accidentally if all master data locations have not been documented accurately. Data can be removed accidentally anyway although there exist confirmations which the user must accept manually, so there must be ways to decrease the possibility that data will be lost forever. Instead of removing a data, user can mark it obsolete whereupon the data will exist also in the future and this is a good option especially in cases where a slight amount of change is needed. Moreover, one of the cornerstones of data management is cyclic backups. There are also other reasons why data can be marked as obsolete and kept in archive. For example, existing laws might create legal requirements which are for instance that specific kind of information must be stored for a specified time and this information must be easily found if asked. Often some kinds of warranties have been

guaranteed to products for a certain period of time which also means that necessary information of products must be kept in archives to be able to provide the agreed warranty functions. [4]

It is always important to make a plan when starting to document organization's internal functions. This is due to the fact that everything is not necessary to be documented. A clear plan makes also the reader easier to recognize why something has been documented, in other words the documented things must be relevant. Requirements of external documentation are often much stricter. For example, product manuals for customers must be very detailed due to existing laws in certain countries. [4]

2.3.4 Evaluating master data

The measurement of data quality level can be expressed as "fitness for use", which means that the value will be formed based on the users' contentment. This emphasizes the importance of knowing the main users of the data, what data these users need and how these users utilize the data. Although certain data would be accurate, it can be inconsistent at the same time leading to situation where the data is unusable. In other words, the quality of data can be proper but the value of data can be slight. The reason for unusable data can be for example that the data can be inconsistently presented or certain necessary information is missing. Creation of cyclic data quality measurements decreases data quality issues which can have for instance negative financial impacts. Data quality measurements are an important part of the master data management. [5]

When evaluating the master data quality, it is important to mind the incidence because often a little number of issues cause most of the challenges. In other words, evaluating the master data quality is worthwhile to begin from the most relevant issues. [2]

Created new master data objects should be accurately evaluated. This means that the object will be checked that it complies with predefined rules which confirm that wrong data will not be put into systems. [3]

3 MDM industrial implementation and concepts

This chapter considers central concepts which belong to MDM implementation in industry. The considered topics are model creation, different implementation styles and change request creation.

3.1 Creation of master data model

Master data management includes a data model creation which will be one of the first steps when establishing master data management. The master data model creation includes four steps [11]:

1. System mapping

2. Metadata collecting

3. Comprehensive planning

4. Maintenance plan creation

The creation of the model starts with system mapping. Mapped systems are either data sources or data destinations. This is an instructive part because the number of different locations might be very much higher than has been supposed. The more there will be different systems the more integrations have to be implemented. It is good to list the source systems and destination systems to be able to specify the scope of the master data management project. [11]

When the systems are known, then the metadata has to be collected. This means for example that there are probably different kinds of data types so there must be mapped how data in systems is related to data in other systems, in other words which are the data dependencies. There are often conversions between data types which can be also related to constraints which define for instance which kinds of values or what size of values can be accepted in each system. [11]

A comprehensive familiarization and planning are important to perform before proceeding into actual master data model developing. The model developing phase precisely defines attributes of the master data records which can be then called the golden records. The master data management model defines how the data will be merged from source systems. The data

will be also cleaned and improved for example with data format normalization, missing value addition, value standardization, attribute mapping and duplicate elimination. [11]

The final phase of the master data model creation is to create a maintenance plan. The objective is to keep the data quality as good as possible. This often means a process which includes designated people which use specific tools to achieve the objective. These people have a knowledge of mastering the master data in its entirety whereupon incorrect data can be recognized straight and the issues can be fixed reliably. Master data model must be modified if related functions change so there must be also designated people who are responsible to organize implementations when needed. [11]

3.2   Implementation styles

There exist different kinds of MDM implementation styles based on the organization's needs. It is important to find out all systems which have necessary information for the MDM development process. Then the purpose of implementing an MDM system must be known. The purpose can be for example that the organization needs a system which is used to create a golden record which is then copied into other systems or the purpose can be for instance that the organization needs a system which gathers and combines information from necessary systems and creates a golden record from them and copies it ahead. Commonly MDM system implementation styles can be divided into four approaches [10]:

- Registry style

- Consolidation style

- Co-existence style

- Transaction/centralized style

The first one is a registry style (figure 2.). It means that the master data is created in other systems and then it is copied into MDM system where it is modified automatically as a golden record. Then the golden record can be used straight in other systems. The registry style is commonly used in organizations where exist a lot of source systems. This style is the cheapest one to implement and the data is only read from the source systems. [10]

Authoring

Read

MDM

Source
systems

Downstream
applications
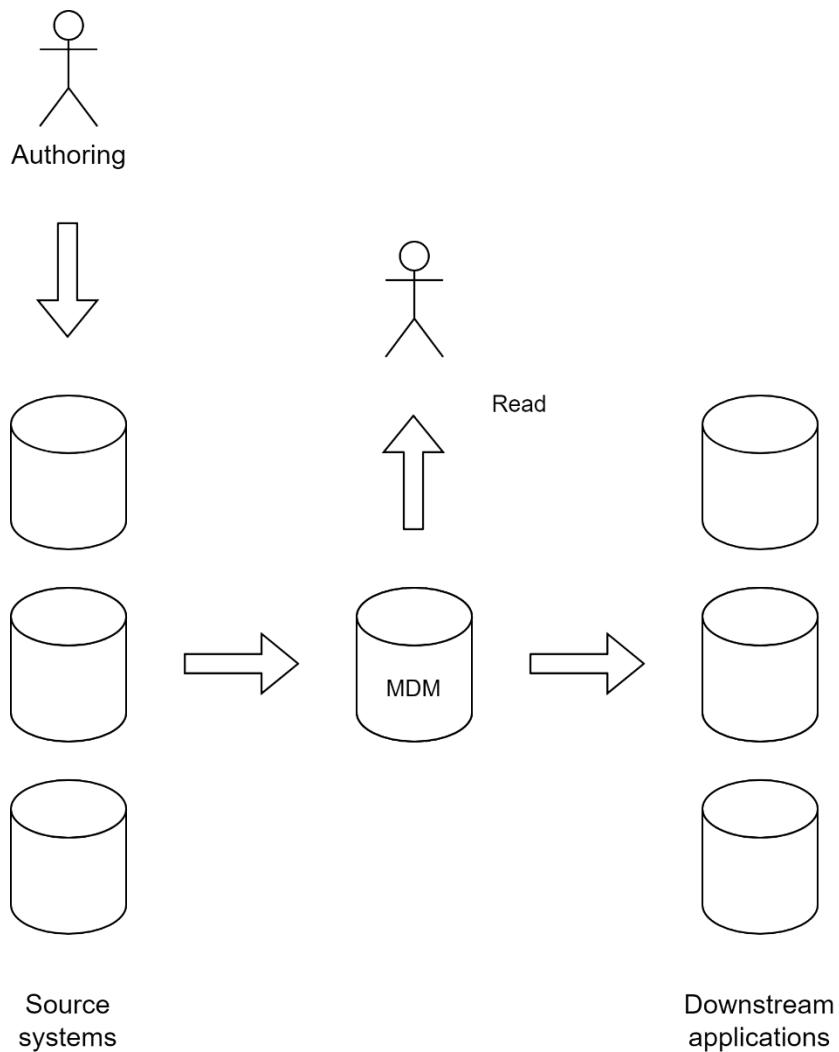
Figure 2. Registry style.

The second way to implement an MDM system is with a consolidation style (figure 3.). This style acts like the registry style but there has been added a human element which makes the automatic golden record process even better with manual activities. In other words, the human is a steward of the master data. [10]
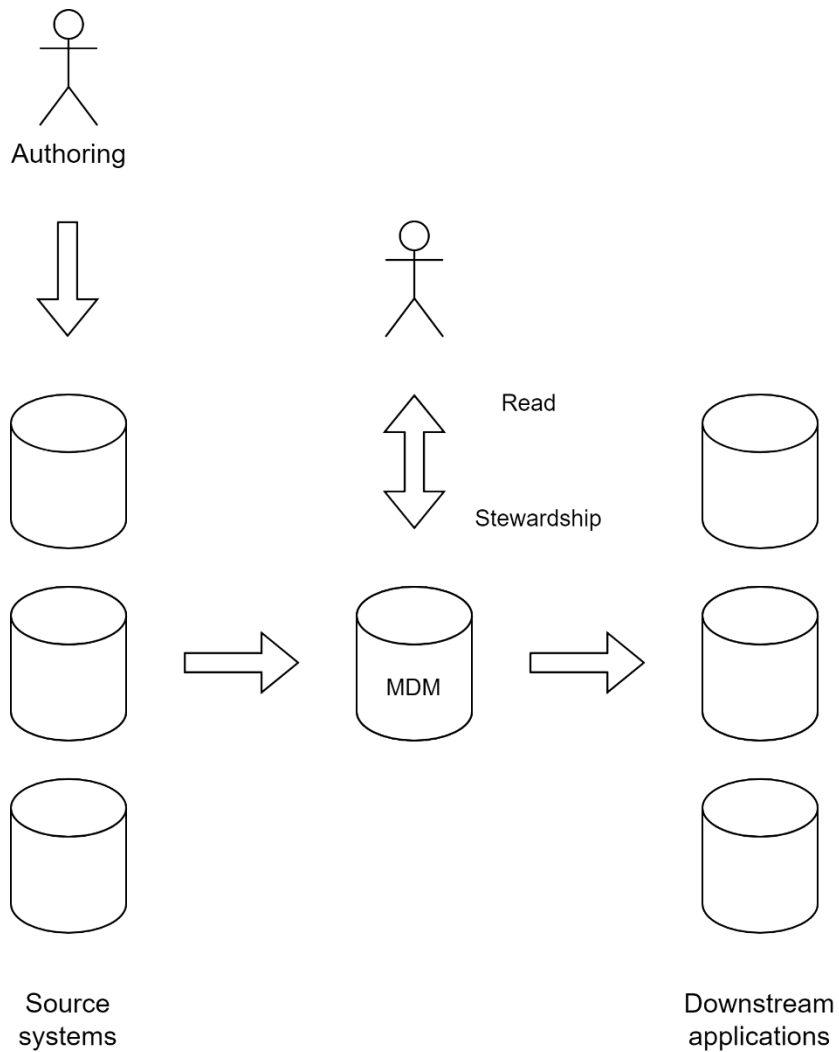
Figure 3. Consolidation style.

The third way is called a co-existence style (figure 4.). This style acts like the consolidation style but the golden record can be copied back to the source systems. This means that the data can be modified in both systems and the modification will be synchronized between the MDM system and the source systems. It is beneficial to notice that the data is not anymore only read from the source systems but the data is also written into the source systems. [10]
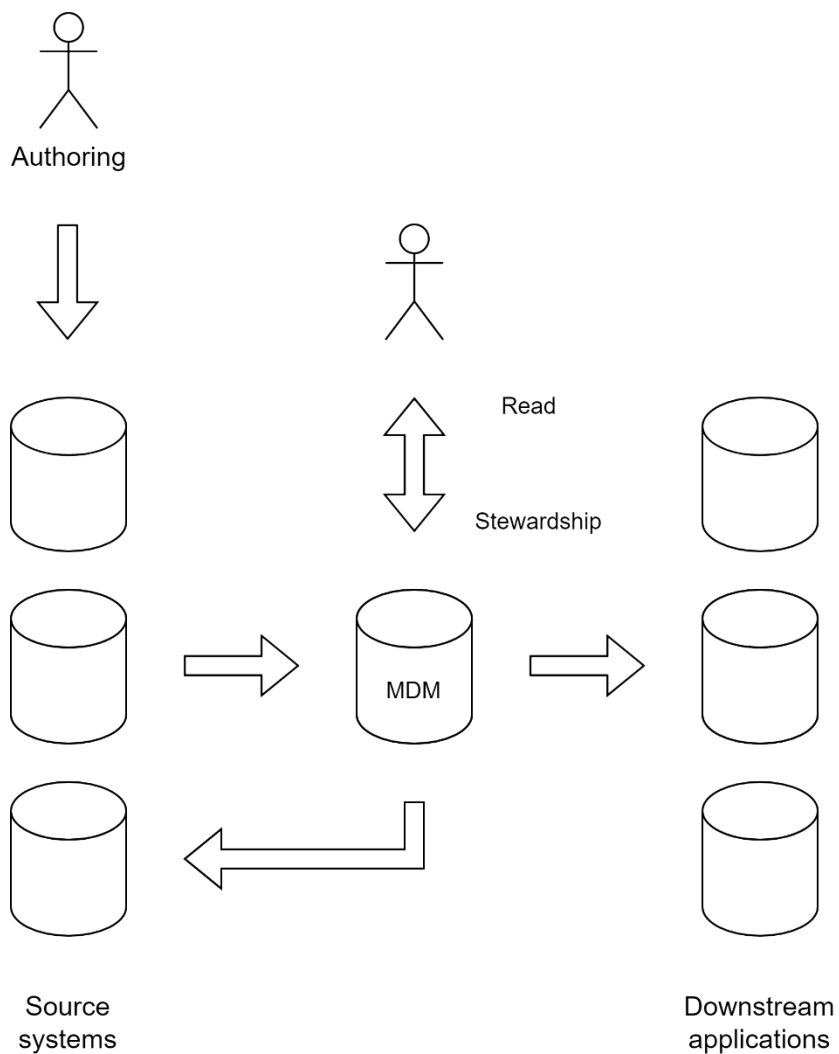
Figure 4. Co-existence style.

The fourth one is a transaction/centralized style (figure 5.). In this approach the whole MDM process starts from the MDM system, in other words all data must be created in MDM system. The other systems only receive the golden record which has been created in the MDM system. In some cases, it can be clearer that the data will be created in one place but this kind of system can take a lot of resources to implement. [10]
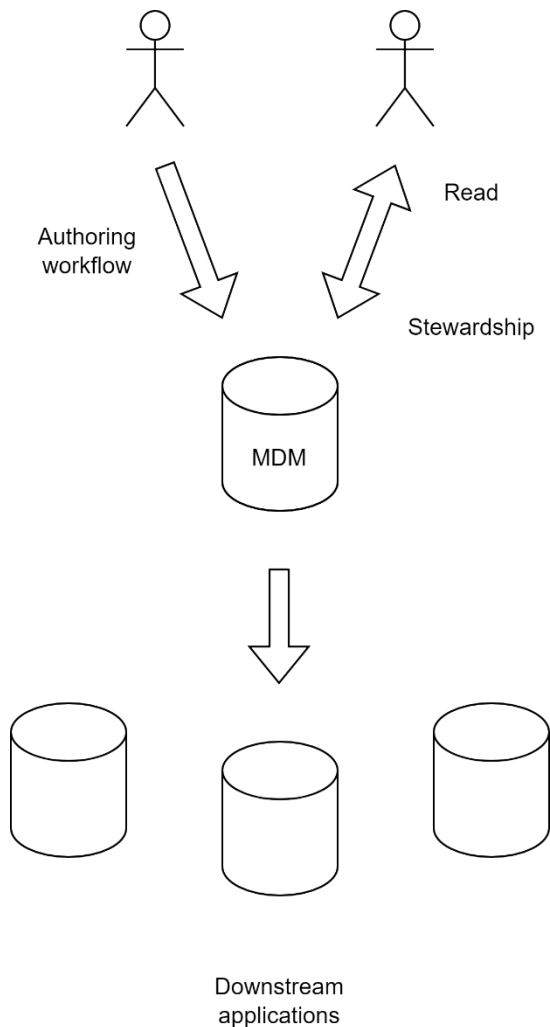
Figure 5. Transaction/centralized style.

3.3 Modification of master data

The changes of master data can be done in a same way as data changes are done normally in enterprise resource planning systems. The changes of master data are beneficial to do with requests which attributes have been defined in the master data change management process definition. The change requests include adding, editing or removing master data. Often the

senders and the recipients have predefined meaning that only certain persons can request master data modification and certain persons can handle the requests. Each request can be divided into two parts. [8, 13]

The first part includes the master data modifications. These modifications can be expressed for instance with some kind of table which includes the exact master data. Often there have been told the type of modification which can be for example new, edit or remove. In case of edit request, it is often required that there have been expressed also the original values that have been modified. In case of remove request, the original values will be expressed naturally in order that the recipient of the request will know which values have to be removed. [8, 13]

The second part includes metadata attributes of the modifications. This gives the recipient a quick overview of the request. This means at least six attributes [8, 13], which are beneficial to be included in addition to the exact master data:

- Request number

- Request date

- Sender

- Urgency

- Target business unit

- Modification description

The first metadata attribute is a request number. This means for example a sequential number which acts as a unique identification for the request. With this information one can refer to some specific change request. [8, 13]

The second metadata attribute is the request date. It can be calculated how much time it took to handle the request when the request date is known. Gathering enough this kind of information, the average lead times can be also calculated. [8, 13]

The third metadata attribute is a sender of the request. This is often the name of the person. With this information the handler of the request can contact right person if something has to be asked. [8, 13]

The fourth metadata attribute is an urgency of the request. This can be implemented for example with numbers from one to three or with literally where number one means low priority, number two means normal priority and number three means high priority. In addition, there could be a description field where the reason for the urgency could be typed because otherwise there might occur a situation where almost every request will be defined as high priority. The urgency attribute tells the handler if the request must be handled as soon as possible or within specified timespan. It could be also beneficial to add a functionality which for instance will send an email to handler if a high priority request has been received. [8, 13]

The fifth metadata attribute is a target business unit. This tells the handler which processes must be followed when handling the request. [8, 13]

The sixth metadata attribute is the modification description. This is a summary of which kind of changes have been made. [8, 13]

After the request has been made, there could be some kind of process status view. This view would tell for example if the handling of the request is in progress or if the request has been handled successfully. It would be beneficial that the view would also include fields which tell who has changed the status of the request and when this has happened. [8, 13]

4　MDM challenges in the case company

This chapter starts with considering what kind of MDM challenges created the need to start planning a new MDM solution for the case company. The reasons and challenges are tabulated in the end of the first section. The second section considers how much there were similarities between common challenges mentioned in the literature review and the challenges faced in the case company.

4.1　MDM challenges and reasons

The investigation of MDM challenges in the case company was started with a review of current MDM process. This included examining how a new data was created in current master data management process and the associated systems.

The current MDM process started with creating master data which purpose was to be transferred into other systems. This data was created within an excel file which was downloaded from company's intranet. With the file one can create new master data or edit existing master data. The existing data was gathered from database into Excel file or it existed in the downloaded file. After new or modified data was created, the file was saved first into the creator's computer and sometimes also into intranet. Then the file was sent to mail group which recipients transferred the data into specified systems. This was a big file so it congested the mail servers.

The mail group received regularly a lot of emails which file content was inspected often by too many recipients of the mail group. This took unnecessary human work because different people were in charge of transferring different data into different systems but the emails did not always have exact information about which kind of data the excel file had.

There was not a centralized way to mark which information was already transferred successfully so some information might be forgotten to be transferred. The creator of the data had to use a lot of time to verify which information was transferred successfully by either asking it from employees or looking it in the different systems. The Excel file did not mark the changed information automatically, so some changed information might be forgotten to mark manually as changed.

The Excel file had some static information which was only maintained in this file so after changing static information the file had to be replaced with the new file in the intranet. This

kind of information was for example information about which products had fixed sales prices although the actual price value was maintained in specified systems. However, the user of the Excel file cannot be sure that the list in the Excel file is always up-to-date. This makes more difficult a situation where for instance not fixed currency prices had to be updated with new currency rates whereupon there will occur a risk where the update process contains fixed prices accidentally if all update targets are not checked carefully. If this risk scenario happens, it will mean that some fixed currency prices will have wrong prices which might lead to situation where some products will be sold with wrong prices.

The Excel file also included currency rates which were stored statically in the file so the creator of new or edited product data must always make sure that the template file had the newest currency rates with which the currency prices will be calculated based on the euro prices. This meant that there will occur a risk where old currency rates can be used accidentally when updating some products' sales prices.

All transfers from Excel file to specified systems took a lot of manual human work for no reason. The challenge of transfers was that the data was created in Excel file but not via some kind of interface straight into the database where the data would be easy to read with automatic software which will modify the data into form which can be imported into specified systems. In addition, some parts of the process contained parts where the data was manually modified in the Excel file before copying into specified systems which might cause mistakes which lead to false information although the original Excel file contained correct information.

While investigating MDM process, causes of challenges and resulted challenges were listed. There was found four main causes of challenges for resulted challenges in the MDM process (table 4.).

Table 4. Reasons and resulted challenges in case company's MDM process.

Causes of challenges	Resulted challenges
Master data modifications were done with downloaded Excel file in the beginning	Using old file can lead for example to wrong sales prices due to outdated currency rates in the file Information can be accidentally not marked as changed in the file whereupon changes will not be done into the destination systems
Master data modifications were sent via email	Mail server congestion Unnecessary file investigation
Updating information into different places manually	Manual interphases increase probability of mistake incidence
MDM process does not have process status view	Investigating the status of sent modification request takes a lot of time

4.2 Challenge similarities between literature review and the case company

The case company had faced same kind of challenges with master data management like has been told in literature review earlier. The analysis compared faced challenges with main challenge types introduced in table 2. The case company's challenges were able to be counted into two main challenge types.

The first challenge similarity was related on both wrong data and insufficient system design. The case company's master data management did not reliably maintain an information about which products had fixed sales prices even though some Excel files can be found which contained this information but the user of these files cannot be sure that the files were always up-to-date. Updating product prices due to for instance currency changes might lead to situation where some products with fixed sales prices can be updated with new currency rates in mass price update although these product prices were not intended to be updated. Wrong sales prices can lead to selling products to customers with wrong prices.

The second challenge similarity was related to broken data. Creating master data in Excel file and sending forward might lead to situation where for example free-text fields had information which was not in correct form due to that there was not checking functionality

which cannot be bypassed. This makes possible to create broken data which other systems cannot accept.

5 New MDM solution for the case company

This chapter considers the new MDM solution which was implemented for the case company. The first section lists the requirements which were specified for the new solution. The second section considers which kind of solution was implemented to fulfill the requirements which were listed in the first section, and which kind of technologies were used in the implementation. The third section considers how the deployment was done. The last section considers the validation of the new solution mirrored to the requirements listed in the first section.

5.1 Requirement specification

The improvement was started with a requirement specification of new system based on the current process information. The main objective was that the new part of the MDM system acts as a single source of truth, in other words the whole MDM process will be started in one place. The specification contained seven main requirements which have been expressed as own sections.

5.1.1 One starting point for data creation

The first requirement was that the new system must be a single source of truth. This meant that all data must be created via one system regardless of whether the data was edited or completely new. The data form must be also better specified which meant that there must be some kind of solution which ensures that the data will be created in a right form. It is important that the data is correct as soon phase as possible so it is worthwhile to spend time to form rules because it will take more time to fix the data subsequently especially if the same incorrect data exists in many systems.

5.1.2 Centralized data management solution

The second requirement was that the new system must be a centralized solution which will maintain data as much as possible so a need to gather data from other systems becomes as little as possible. All data must be stored into database which removes the need for data maintenance in Excel files. The objective was that the data needed in MDM process would be located at the same place.

5.1.3 Process automation

The third requirement was that the master data management process must be improved also with automation. The most important objective was to reduce manual work in the process. The master data management process included steps which require manual work which actual purpose was for example to create files which had wanted information in the form which can be imported into other systems. This information was gathered from the Excel file which was received from master data management mail group. This kind of data copying from one place into second place was always performed similarly which means that could be easy to be replaced with automatic data processing. The key for multiple automatic operations will be that the data is stored into database via interface which makes sure that the data is in proper form.

5.1.4 Easy to access

The fourth requirement was that the system must be easy to access for every employee. This meant that there must not be a lot of things to do before the system's basic functionalities can be accessed, meaning for instance that there must not be a lot of software which have to be installed to be able to do basic MDM operations. The system should be only used with personal account. This is a requirement due to the management of access rights and that it would be easier to contact right person in respect of certain things which have been performed in the system.

5.1.5 Process status view

The fifth requirement was that the status of data transfer process must be visible and easy to consider. This meant that the author of new data must know when the data will be available in every system. It is good to know for example when a necessary product information will be available in sales system in order that for instance salespersons can be informed when certain product information is available. This kind of view must be located in one place where the progression of whole data transfer between systems can be inspected. The view must show the status of requests in different systems separately with related necessary metadata.

5.1.6 Gathering supplementary quantitative information

The sixth requirement was that the system must gather metadata and usage log into database. This information will act as a supplementary quantitative information. Timestamps produce

information which can be used to calculate time information of data transfer in respect of master data management. With the statistics one can for example improve lead times by seeking bottlenecks of data transfer through systems. Gathering usage logs reveals for instance how much new products have been opened via the system or when a certain product information has been modified via the system.

5.1.7 Easy management of access rights and basic maintenance functions

The seventh requirement was that the system must have access rights management and basic maintenance management functionalities via graphical user interface. With the access rights management functionality, specific users or user groups can be given rights to specific functionalities or views in the system, in other words different employees will have different roles in the master data management process which will make the process more stable when all employees cannot do the same things. It must be possible to give same permissions as a group to employees which for example have same official role in the company but permissions must be able to be given also to specific users as well. Allocated permissions will make sure that the employee with certain permission is able to do certain part of the process correctly due to that the employee might for instance have a proper education in respect of the work assignment. In addition to access rights management, the new system must include functions with which basic maintenance can be done including specifying mail groups, mail server address and product price update related settings.

5.2 Implementation

Planning of the new system was started with analyzing the requirement specification information. Existing systems in respect of master data management process was mapped to be able to know how the whole improved master data management process will function. The master data system did use transaction/centralized implementation style with a certain addition (figure 6.). This addition was that the MDM system did use downstream applications' master data in editing functionality but the data is same that has been created earlier in MDM system. The addition has been visualized with an additional arrow from downstream applications to MDM in the figure. The reason for using modified implementation style was that this project's scope did not include a part where the MDM system would show existing master data based on the status of each master data modification request. An example of this situation is that the MDM system should show the same product

sales prices which salespersons see, in other words the new sales prices must have been gone through in the end of the MDM process in order that salespersons see them whereupon the new sales prices can be used also in the MDM system.
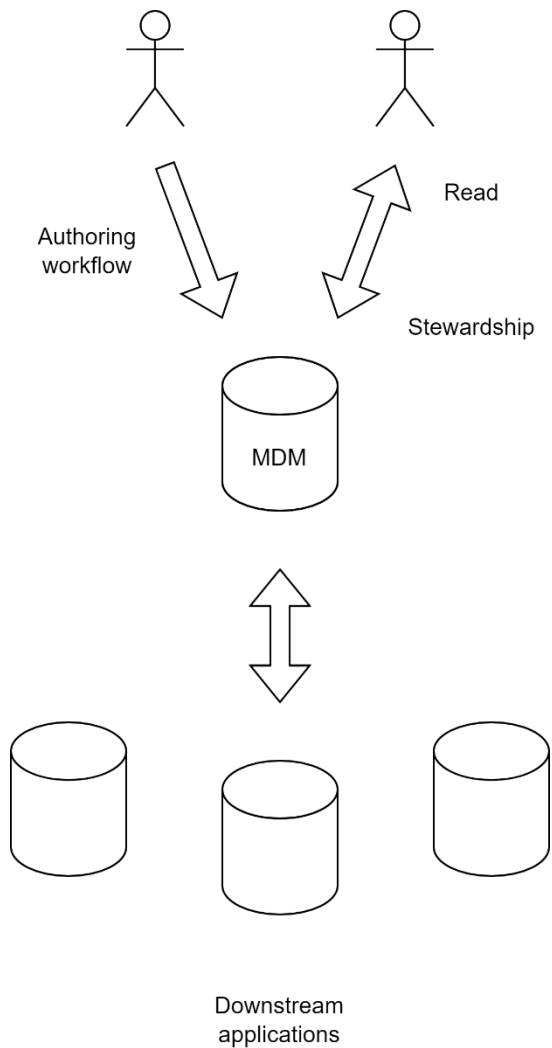

Figure 6. MDM implementation style in the case company.

5.2.1 Meeting requirements

This part considers how the earlier specified seven requirements were fulfilled. Each requirement has been discussed separately in own section.

One starting point for data creation

The one starting point for data creation -requirement can be fulfilled technically with a web application because the process can be improved with a way where the application's inputs are the only acceptable data sources. The application can also take care of data validity with predefined rules. In order that the requirement can be fulfilled entirely, also the master data creation process had to be modified.

Centralized data management solution

The centralized data management solution -requirement was solved next. The system must have a database which must have necessary data as much as possible. The original process only gathered part of the data from the database so the new database was created into the same database server due to centralizing same kind of data into same place. Gathering all master data into one place also made the creation of golden records easier. This solution also made reporting more reliable.

Process automation

A key for process automation -requirement's fulfilling was that more data will be available in database in the beginning of the process. For example, it is easier to copy data forward programmatically from database than gather data into Excel file and then process it forward manually. This requirement was fulfilled with a way where all data must be stored into the database with predefined rules whereupon the data can be processed automatically between different systems. Automatic data processing also will decrease the amounts of mistakes caused by manual data processing.

Easy to access

The easy to access -requirement was the starting point for meeting the requirements. The new system must be easy to access for every employee. All necessary employees had access to company's own internal network so a web application was a natural solution to meet the requirement. In other words, all what the user needed to access the system was a web browser. In addition, the system did not require separate login-page due to that the access control was implemented with Active Directory -functionality. This made the using of the new system easier because the user was allowed to access the system if the user was logged into the company's internal network.

Process status view

The process status view -requirement can be implemented with own page in the web application. This page will show all new data modification cases as a table. This process status view table will show the status of data copying process through systems. There were used six metadata attributes:

- Request number

- Request type (new/edit)

- Sender of the request

- Request date

- Transfer metadata of specific downstream system

- Modification description

The table only contained columns which were necessary to be able to see quickly and easily the status of for example own master data modification requests. The status of each step was presented with own background color in addition to the status text in order that it would be easier for human to perceive the situation when for instance green background color of the field indicates that the processing was done successfully. The process status view included filters with which one can see requests which are for example in progress.

Gathering supplementary quantitative information

The gathering supplementary quantitative information -requirement can be implemented with storing metadata and usage logs into database. Especially it was important to gather statistical data which was related to the master data processing between systems in order that the data was able to be used to find bottlenecks in data transfer process. This gathered data was used for instance with business intelligence -tools which revealed that certain parts of the master data processing between systems were slower than other parts.

Easy management of access rights and basic maintenance functions

The easy management of access rights and basic maintenance functions -requirement can be implemented with own page in the web application in order that users can maintain functionalities with graphical user interface. This makes the maintenance easier for employees

who are not information technology professionals. The page is only available for predefined employees. The page contained two entities which were access rights management and basic maintenance functions. The access rights management was implemented with a way where specific rights can be given to users based on Active Directory groups or specific users. The basic maintenance functions included for example specification of mail groups, mail server address and product price update related settings.

5.2.2 Technologies

The meeting requirements -section gave background information for some technology selections. These were for instance that the new master data management system will have a database and the usage interface will be implemented with a web application.

The case company uses a lot of technologies by Microsoft PLC so the choices were emphasized to technologies which are commonly known in the company. The development was easier to start when for example there was less need for installing applications and development environments. In addition, with familiar technologies other employees can develop extensions easier to the system.

The case company had a lot of Microsoft Windows servers which had IIS web server softwares and ASP.NET frameworks installed, so it was a natural choice to use these. The ASP.NET framework can be used for instance with a C# programming language which was chosen as a backend programming language for the web application due to personal experience. The Windows servers had also Microsoft SQL Servers installed so it was also a natural choice to use these. The web application's user interface was implemented with HTML, JavaScript and CSS languages.

5.3 Deployment

The deployment contained two parts. The first part considered the planning of the deployment. The second part considered the actual deployment.

5.3.1 Deployment planning

The deployment planning was started with defining how the transition from using old system to using new system must be done and what must be taken into account. This meant that the deployment must be thought both from technical and employee point of views.

The technical deployment can be done beforehand because the new system can be available in the corporation's internal network but access control must be used to make sure that only wanted persons can access the system in testing phase. This testing phase was divided into six steps (figure 7.). These steps formed four entities. The first entity contained testing that the system is available in corporation's network with limited access. The second entity contained testing that existing real master data can be gathered into the new system and regular data updates work correctly. The third entity contained testing that creation and editing of test master data and forming this data into other systems' accepted forms will work correctly although the test data will not be transferred into other systems. The created test master data must be removed before the last testing entity will be started. This last entity included testing the new system with real master data. This included the creation of new master data and editing existing master data functionalities and forwarding this data through the whole process. The real master data testing was naturally implemented by real end users which consisted of a couple of employees.
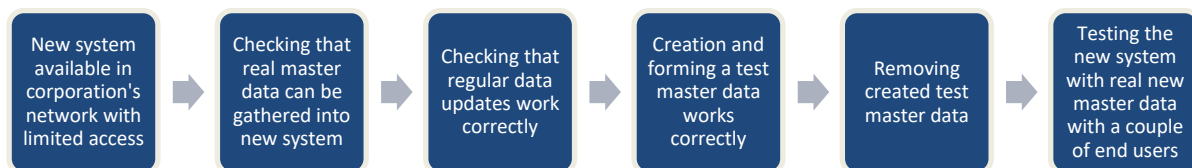

Figure 7. Steps of technical deployment.

From employee point of view, the transition to using new system must be done with a way where existing unfinished master data modification requests must be finished first. This means two things. The first thing is that users finalize their existing master data modifications and send them forward with old way in order that unfinished not sent requests do not have to be done again in the new system. The second thing is that the requests that have been sent forward but have not been gone through to destinations must be done ready. In addition, the new system must be in place for new requests if users have not started to create new master data in their computers.

The transition to using new system requires that a training must be organized for users that they know the new process and the related technical system. There will be three trainings which all will include same steps (figure 8.) but with a little bit different content. The first

training will be organized for users which will maintain the system. These employees are working physically at the parent company of the case corporation. The second training will be organized to employees which will work with the system and are working physically at the parent company. The third training will be organized to employees which will work with the system and are working physically at foreign subsidiaries. All three trainings must contain a part where has been told that how the transition must be done from using old system to using new system. In beginning of the first training, the new system will be available to the maintenance group whom the first training will be organized. In the beginning of the second training the new system will become available to all employees but a part of the second and third group's employees will not have access to all functionalities.

Figure 8. Steps of MDM training.

5.3.2 Deployment

This part has been divided into two subparts. The first subpart considers how the new MDM solution was deployed technically and how it was tested. The second subpart considers how users were trained.

Technical deployment and testing

The technical deployment was started with ensuring that the system in production environment was accessible only for allowed users defined in the access control list. This contained testing that the access control both with Active Directory groups and individual users worked as planned. This was ensured with making tests with employees who were defined in the access control functionality. The access rights for these tests were given only for the time period of testing. After successful testing the access rights were removed.

The second testing phase included testing that real master data can be gathered into the system to be able to make modification requests of existing master data. This was handled with one functionality at a time. The comparison was made between the master data in database and the master data which the new system expressed in the user interface. The master data in database was remained constant during this testing phase.

The master data was expressed correctly in the new system whereupon it was time to test the regular data updates. These data updates were implemented with automatic scheduled tasks. The new data in database was also gathered into other tables whereupon it was easier to ensure that the new data was expressed correctly in the new system's user interface.

The next testing phase included creation of new test master data and editing existing master data. There were different use cases which were documented during the project. These use cases were implemented with a way which the data will be wanted to be used with the new system. This testing was done forward until the next step would have been to transfer the created or edited master data into next systems. The last step was that the created or edited master data was converted into forms which the other systems will accept. This testing phase was not implemented entirely in the production mode due to that the automatic notifications of new or edited master data were not sent to the real mail groups which the new system will use when the testing has been done but the notification mails were sent only to the creator of the master data.

Earlier phases included creation and editing test master data which was stored into different locations whereupon the new system contained data which might confuse real end users when the new system will be in place entirely. To tackle this, the test master data, which was stored by the new system, was removed with a way where the used databases were wiped completely and necessary existing real master data was gathered into the new system again.

The testing was moved forward into phase where the new system was stated to work properly with test master data. The last testing phase was done with real new master data which was created by real end users. To enable this, necessary access rights were given to a couple of employees for the time period of last testing phase. The testing included creation of totally new master data and editing of existing master data. This data was forwarded through whole process including transmission into other systems and gathering this data into the new system to be able to see that the data is available for next possible editing cases. This testing phase was neither implemented entirely in the production mode due to that the automatic

notifications of new or edited master data were not sent to the real mail groups which the new system will use when the testing has been done but the notification mails were sent only to the creator of the master data. After the test was completed successfully, the given access rights were removed.

User training

The user training was started with surveying who will be the employees which will work with the master data management and need to use the new system. In addition, maintenance group of the master data management was surveyed. The notification mails' recipients were formed from the surveyed employees. These mails are important for example in the improvement of master data's lead times. Based on the surveyed employees, it was also possible to form access control groups into the Active Directory but these groups were not yet selected in the new system.

The earlier surveyed employees were divided into three groups which formed the target audiences of the three user trainings which will be organized. The first group included employees who will maintain the new system. All of these employees were working physically at the parent company of the case corporation and the training was organized face-to-face at the headquarters. Part of this group's members were along specifying requirements for the new solution and testing the new system with a real new master data. The second and the third groups included employees which were end users of the new system but they not maintained it. The difference between the second and the third group was the way with which the trainings were organized. The second group's members were working physically at the parent company of the case corporation whereupon this training was also organized face-to-face at the headquarters like the first training. The third group's members were working physically at foreign subsidiaries whereupon this training was organized with online video meeting.

Before the first training, the new system was put available with correct access rights for the employees of the training and before the second training the new system was put available with limited rights to all employees of the case corporation. This made possible for employees to start familiarize themselves with the new system and possible questions in respect of its functions can be asked immediately during the training.

All three trainings started with a review of the old way to manage master data. This contained the old process and the old technical tools. During this review, earlier faced challenges were told when they came into sight.

Then the new master data management process was introduced. The new process was gone through step by step. The advantages of the new process were told in addition to the ways with which the earlier challenges have been resolved. All differences between the old and the new process were considered meticulously.

The new technical solution was introduced next. This started with going through how the new system must be used to comply with the new process. Like the earlier process comparison, this section contained also part where the similarities between the old and the new system were expressed. This comparison helps people to perceive the transition better related to for instance where certain needed things can be found from the new system's user interface.

All trainings told the system maintenance functionalities but the using of these was trained more deeply for the maintenance group. For the audience of the second and the third trainings were told that they have to contact the maintenance group if some system maintenance related work must be done and it was also told that how the contact information of this maintenance group can be found.

The last part of the trainings explained how the transition to using the new system must be done. This started with telling that employees can naturally finish all master data modifications which are in progress in employees' computers. It was emphasized that these must not have to do again in the new system but the transition will be done as smoothly as possible and with as few auxiliary work as possible. However, after the transition period the master data modifications are not allowed to do with the old way but these must be done with the new way. So, the using of new system must be started immediately when possible.

After trainings, a notification about the new way to manage master data was published in corporation's intranet. In addition, foremen were asked to inform employees who were invited to trainings but were prevented to attend.

5.4   Validating the solution

This section considers how the specified seven requirements have been validated as resolved. This has been presented as a table which each row considers a validation of one requirement (table 5).

Table 5. Validating the solution.

Requirement	Validation
One starting point for data creation	This requirement was fulfilled with process improvement. This meant that an only allowed way to create new master data was via the new web application where the data is stored straight into database. This new way was trained for end users with trainings where was highlighted that other ways to create a new master data are no longer allowed. This validated that there is only one starting point for data creation.
Centralized data management solution	This requirement was fulfilled with a new data architecture. Data is no longer maintained or gathered from Excel files but all information is stored into one database where the golden records will be made. This validated that the master data is managed as centralized.
Process automation	This requirement was fulfilled with creating automatic data processing. This decreased manual interphases which validated the solution for the requirement.
Easy to access	This requirement was fulfilled with a web application which is available to all employees in the corporation. The application is running in the corporation's internal network whereupon it was also easy to implement access control without separate login-page because the access control was made with Active Directory -functionality. The testing also indicated that everyone was able to access the web application via web browser. This validated that the new solution was easy to access.
Process status view	This requirement was fulfilled with own page in the web application. This page showed all required metadata about master data modification requests including filters which can be used to examine specific results. The view also included buttons with which each request's master data modification can be seen. These validated that the solution was implemented as specified.

Requirement	Validation
Gathering supplementary quantitative information	This requirement was fulfilled with storing metadata into database during the master data transfer between systems. This metadata had necessary statistical information in order that it can be used for instance with business intelligence -tools to find bottlenecks in data transfer. This validated that the system gathered necessary supplementary quantitative information.
Easy management of access rights and basic maintenance functions	This requirement was fulfilled with own page in the web application in order that maintenance group can handle maintenance functionalities with graphical user interface. The page included access rights management which was implemented with user friendly lists where user can choose groups or users which rights are wanted to be changed. The mail groups, mail server address specification in the page has been implemented with a way where typed settings are validated automatically whereupon the application works as planned. The product price update setting has been implemented in a page with a way where user can easily choose whether the update setting is either on or off. These validated that the access rights and basic maintenance functions can be easily managed.

6   Conclusions

The goal of this thesis was to research master data management in theoretical context and then implement a new master data management solution for the case company. The first section considers the key findings in theoretical part, the selection was based on which topics were most emphasized in literature review. The second section answers the research questions.

6.1   Key findings

The value of well managed master data has been noticed increasingly in companies. It decreases unnecessary expenses but also helps to make better decisions.

Master data management is not only an information technology project because it integrates business and information technology operations. It differentiates from normal data mastery policies due to its company-wide scope.

The business needs are usually changing which means that also master data management needs to change to be able to respond to changed business needs. The changes will have different dimensions depending on which layer the change will have effect. The changed policies can exist either in strategic, tactical or operational layer.

6.2   Answers to research questions

The thesis had three research questions. Each question has been answered in own part.

6.2.1   What must be taken into account when implementing master data management in industry?

The answer to this question includes three steps which are found to be the most important ones. The steps comply with chronological order of MDM implementation.

The first step is to plan the master data model as well as possible. This acts as a base for other steps.

The second step is to choose the implementation style carefully. Different organizations have different needs for master data management so the implementation style must be chosen mirrored to the needs. These needs can change whereupon it is worth to check if the current implementation style is the best still or should it be changed also.

The third step is that master data change requests are worth to be planned well before starting to use the master data management. The attributes of each request must satisfy at least the needs of whole master data management process.

6.2.2 How the current master data management process in the case company can be improved to fulfill predefined requirements?

The process improvements were implemented mainly to the beginning of the old master data management process which was enough to fulfill predefined requirements. This meant that the new process allowed a new master data to be created only via the web application which was easy to access for all employees. The new process decreased manual interphases due to increased automatic data processing.

6.2.3 What kind of technical solution the new master data management process will need in the case company?

The implemented technical solution was a web application which was available to all employees in corporation's internal network without separate login. The creation of new master data must be done with predefined rules which have been defined in the application. The editing of existing master data has been done with golden records which have been formed in the centralized database. The master data has been processed more automatically whereupon it is easier to transfer into downstream systems, in other words with less manual interphases. The web application's user interface has a process status view which showed the status of each master data transfer process. A metadata will be stored into database during the processes and this data will be used for example in process status view or with business intelligence tools. The maintenance functionalities were implemented with own page where maintenance group's members can adjust settings. The access control has been done with Active Directory -functionality.

References

[1] Vilminko-Heikkinen R. (2017). Data, Technology, and People: Demystifying Master Data Management. Doctor of Philosophy Thesis, Tampere University of Technology.

[2] Heiskanen M. (2016). Data Quality in a Hybrid MDM Hub. Master of Science Thesis, Tampere University of Technology.

[3] Ikola A. (2018). Developing master data management in a multi-business case organization. Master of Science Thesis, Tampere University of Technology.

[4] Kortesmaa N. (2016). Remodeling master data in changing business environment. Master of Science Thesis, Tampere University of Technology.

[5] Pietarinen S. (2017). Increasing the value of good quality customer master data in a global enterprise. Master of Science Thesis, Lappeenranta University of Technology.

[6] Papa A. (2017). Master Data Management in Global Enterprise. Master of Science Thesis, University of Turku.

[7] Scheidl H. (2011). Master data management maturity and technology assessment - From theory to practice, case Ineo Oy. Master of Science Thesis, University of Turku.

[8] Firican G. (2018). Learn how to create a MDM change request. Https://www.lightsondata.com/mdm-change-request-template/ (referred 5 / 2021).

[9] Cleven A. & Wortmann F. (2010). Uncovering Four Strategies to Approach Master Data Management. In proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS 2010).

[10] Dalimunthe, R. (2019). Understanding Various MDM Implementation Styles. Https://towardsdatascience.com/understanding-various-mdm-implementation-styles-5b4c8fcbbecf (referred 12 / 2020).

[11] Haselden K. & Wolter R. (2018). Master data management - what, why, how & who. Https://profisee.com/master-data-management-what-why-how-who/ (referred 5 / 2021).

[12] Teleste - Company. Https://www.teleste.com/company/ (referred 3 / 2022).

[13] Parhizkar, M. & Comuzzi M. (2016). A framework for impact analysis of post-implementation enterprise resource planning modifications. In proceedings of the Computing Conference (SAI 2016).

[14] Suram S. & Muppala R. (2008). Master Data Management-CDI. In proceedings of the IEEE Region 10 International Conference (TENCON 2008).

[15] Wang L., Ming X. & You J. (2009). The Steps and Methodology of Identifying Master Data from Business Processes. In proceedings of the WRI World Congress on Software Engineering (WCSE 2009).

[16] Karia J., Sundararajan M. & Raghavan G. (2021). Distributed Ledger Systems to Improve Data Synchronization in Enterprise Processes. In proceedings of the IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER 2021).