Turun yliopisto
University of Turku

MASTER DATA MANAGEMENT MATURITY AND TECHNOLOGY ASSESSMENT

From theory to practice, case Ineo Oy

Master´s Thesis
International Management of
Information Technology - IMMIT

Author:
Half Abude Scheidl

Supervisors:
Ph.D. Timo Leino
Ph.D. Paul Laifa
Ph.D. Anna Rutkowski

31.05.2011
Turku

Table of contents

List of figures

List of tables

List of abbreviations

- CIAM – Corporate Information Asset Management
- CRM – Customer Relationship Management
- MD – Master Data
- MDM – Master Data Management
- ERP – Enterprise Resource Management
- IS – Information System
- MPC – Manufacturing Planning and Control
- MRP – Material Resource Planning
- KPI – Key Performance Indicators

1. INTRODUCTION

1.1 Background for the study

The subject of Master Data (MD) and Master Data Management (MDM) is an emerging Information System (IS) research topic which is experiencing a *hype* phenomenon similar to what happened to other technologies such as Enterprise Resource Planning (ERP) and Data Warehouses. At the same time, an increasing number of authors acknowledge the existence of the underlying problems that technology vendors claim to solve (Smith and McKeen 2008; Radcliffe 2009; Ventana Research 2009), such as data inconsistences and key business entities having multiple definitions.

The special attention to MDM is due to increasing number of and complexity in regulatory policies and the unfulfilled promises of other enterprise systems to deliver a single version of key business entities, such as customers and products. As a consequence of mergers and acquisitions common to globalization strategies and market consolidation trends, the actual record of a company's *customer base* might involve data scattered in dozens of systems. Harmonizing those business definitions and data records across business lines and across subsidiaries is no simple task, requiring restructuring of data ownership and governance, however also requiring sophisticated technology for policies and business rules to be enforced.

Thus, it is valuable for enterprises considering implementing MDM to understand the role of technology in solving its data problems, and how to assess when and how a new tool should be integrated to their enterprise systems topology. Building a business case for MDM is said to be the first and most difficult step.

1.2 Research contextualization

This master's thesis was written for International Master in Management of Information Technology (IMMIT), a master degree from the Erasmus Mundus Action 1 program, sponsored by the European Commission for Education and Training. IMMIT is organized by Turku School of Economics (Finland)[1], Tilburg University (The Netherlands) and IAE Aix-

[1] Turku School of Economics (Turun Kauppakorkeakoulu) is since January 2010 part of Turku University (Turun Yliopisto).

en-Provence Graduate School of Management (France). IMMIT master's theses can be connected to an internship, so that research is conducted according to company's needs, the case of this research. This research in particular was carried out for Ineo Oy[2], and supervised by Doctor Timo Leino from Turku School of Economics. Ineo Oy (referred in this thesis as Ineo) is a Finnish privately held IT services and solutions corporation founded in 1999, offering advanced solutions for Corporate Information Asset Management (CIAM). Ineo holds a leading position in ERP data migration in Finland and serves customers in advanced data warehousing and data integration needs.

In 2010 Ineo initiated a program called CIAM Suite, aiming at developing the supporting technology for their CIAM conceptual framework, which evolved from years of expertise in data quality, data migration and master data projects. CIAM Suite is a Research and Development project, partially funded by Tekes[3] *Digital Product Process* program.

The main driver for this research was the need for a comparative perspective on the CIAM Suite project. The main project manager demonstrated interest towards a comprehensive analysis of current concepts and tools on the MDM subject, aiming at an output which could provide insights on the scope of the product being developed, such as software features and system architecture.

1.3 Purpose of the study

This research concentrates on two main areas of interest, with two complementary purposes and outcomes. Firstly, it aims at investigating the current conceptual maturity level in the field of MDM by means of a literature review and primary data collection from interviews. The outcome of this research phase is a framework in which Data Management products or solutions can be evaluated or "benchmarked".

Secondly, this research aims at providing Ineo with practical subsidies that enable its managers to enhance or evaluate the scope of the CIAM development project. Ineo faces the challenge of materializing its tacit expertise in Data Management into a product which reflects its vision. It would be a valuable asset for Ineo to be able to compare it with a comprehensive view on *what* such a product should deliver.

---

[2] http://www.ineo.fi

[3] The Finnish Funding Agency for Technology and Innovation. Tekes is the main public funding organisation for research, development and innovation in Finland. From: http://www.tekes.fi/en/

This research will use a qualitative approach to answer the following research questions:

1. What is the current state of Master Data Management concepts and practices?
2. How a Master Data Management solution can be evaluated?

1.4 Thesis organization

This thesis organized in literature review, research design, research results and conclusions and summary (Figure 1).

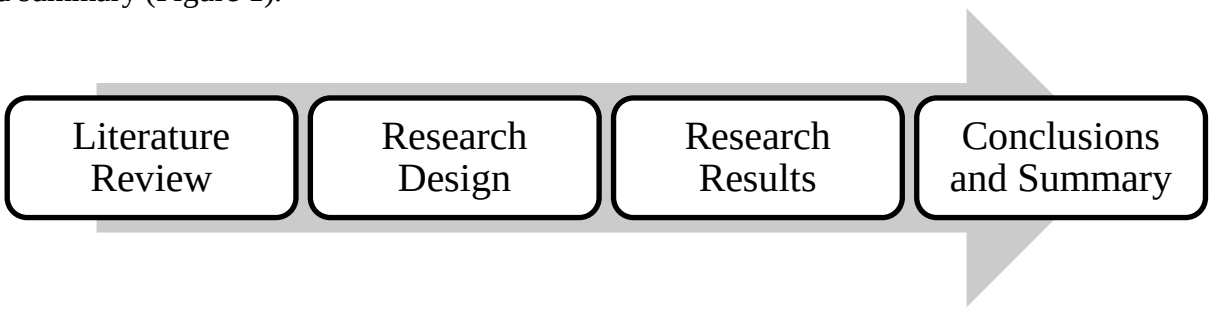

Figure 1 Thesis organization

The literature review is divided into chapters 2 and 2.4, which are aimed at presenting a review of existing literature on the subject of *MD* and *MDM* concepts and practices. The review was organized around the two main research interests groups, namely *what* constitutes master data management and *why* it is a relevant topic in IS (section 2), followed by an examination of *how* MDM should be developed in a corporate environment (section 2.4). The "what" and "why" aspects were included in order to present a comprehensive view on the subject, based on the researcher's assumption that MDM is an emerging topic, and are also briefly covered in the empirical part. However, the "how" part is the most relevant in the context of this master thesis, and is also emphasized in the empirical part.

In this research, the approach adopted for Literature Review was a "systematic, explicit and reproducible method for identifying, evaluating and synthesizing the existing body of completed and recorded work produced by researcher, scholars and practitioners" (Fink 2005). The adopted classification or taxonomy for this literature review according to (Cooper 1988) is presented at the beginning of section 2.

Section 4, Research Design, details how the research process was organized and the underlying theoretical models for the adopted research method (Qualitative Research). The steps for data collection are explained along with a description of the participants in the

study, in terms of their experience and role. In Section 6, research results are presented and analysed.

The last section, Conclusions, highlights the research implications of this study, and proposes particular areas for further investigation. This section also summarizes the practical implications for the sponsor of this research, with the selected suggestions taking into consideration the demand for actionable, tangible conclusions from the part of the organization.

2 THE WHAT AND WHY OF MASTER DATA MANAGEMENT

The purpose of this review is to “understand the structure of the subject”, “identify relationships between ideas and practices”, “relate ideas and theory to application” (Randolph 2009). According to Cooper’s taxonomy of literature reviews (Randolph 2009), this research is mainly focused on research outcomes and on practices and applications of MDM, with the goal of identifying central issues. Its coverage is a representative sample of articles and books with an “espousal of position” perspective, which suits better the qualitative nature of this research (as opposed to a “neutral representation" traditionally found in quantitative research). The content is organized around main concepts, as opposed to historical or methodological organization). Finally, the audience comprises scholars and practitioners.

Table 2-1 Literature review taxonomy of this research (according to Cooper's framework)

Characteristics	**Category**
Focus	Research outcomes Practices or applications
Goal	Identification of central issues
Perspective	Espousal of position
Coverage	Representative
Organization	Conceptual
Audience	Specialized scholars Practitioners

2.1 MDM as an emerging research field

A common issue with emerging technologies such as Master Data Management is the lack of critical mass from academic research. There are few articles on the subject from journals such as *MIS Quarterly* and *Communications from ACM.* As an example, performing a query on the subject “Master Data Management” on ACM Digital Library yields merely 94 results, while the same search term on Google yields around one million results[4].

Not only few references can be found on real implementation and business value of MDM references, but also the definitions are broad, having authors emphasizing one particular

[4] Results from March 15th, 2011. As a comparison, the search term "enterprise resource planning” yields six million results on Google and around 2000 on ACM Digital Library.

aspect or another, such as workflows to enhance governance practices or data integration to make MD available to other line of business applications. (Wadehra 2007) highlighted the diverse requirements have not yet converged into a consistent market, thus making companies confused on where to start.

In addition, it is said that MDM is more of a technology being pushed by vendors than a solution for a real problem. (Sammon, Adam et al. 2010) argue that MDM is effectively "Data Warehousing branded with ERP market rhetoric and contains an added repository of 'master data'". The authors emphasize the need for research in order to prevent similar unfulfilled promises of earlier Data Warehouse and ERP projects.

(Silvola, Jaaskelainen et al. 2011) also reinforce the lack of a common understanding on what MDM is, and conducted an interview study in eight high-tech companies, proposing "How to define one master data?" as one of three research questions. (Smith and McKeen 2008) conducted a focus group aiming at delineating MD boundaries and concluded the "definition of MDM is vague".

On the other hand, publications from vendors and research companies which do not relate to universities are abundant. Usually referred to as "research papers" and "white papers", these publications do not elaborate on the research methodology or sources of data, imposing an additional challenge on selecting authoritative references for academic research aimed at emerging IS applications.

Nevertheless, research companies have shown steady growth and profitability, and not rarely including in their shareholders and board of directors members from the same players that are object of their research, a clear reason for criticism on the independence of these research institutions and how much vendors influence the research agenda (Greenemeier and McDougall 2006).

In this thesis, there was a clear objective of contrasting perspectives from academic and commercial related sources. In this thesis we refer to these two groups of references as *scientific literature* (references written by professors and researchers from universities and published in peer reviewed journals for example) and *semi-scientific* literature, published by market-oriented research institutions such as Gartner, Forrester and Ventana Research or the research units from IBM and Microsoft.

2.2 The evolution of data processing and data management

When computer resources became available to enterprises in the 60's, data in digital format was centralized in a few locations, thus eliminating the risk of duplication or redundancy. At that time, a "single computing resource housed all the applications and all associated data sets" (Loshin 2009), and it was fairly easy to maintain a single set of data about the basics of

business, such as product information and customers characteristics (Ahern 2006) . At that time, cost was the primary business competitiveness driver, and manufacturing strategies were highly product-focused, aiming at high volumes and cost minimization, including inventory management (Robert Jacobs and Ted' Weston 2007). The immediate application of electronic data processing capabilities was in inventory management and computerized Reorder Point[5] systems.

> *"Early MRP [Material Requirements Planning] application software was the state-of-the-art method for planning and scheduling materials for complex manufactured products" (Robert Jacobs and Ted' Weston 2007).*

The evolution in microelectronics and programming languages in the 80's made computers and software more powerful, affordable and extended its applications from pure material management to accounting, order entry and invoicing, forecasting, and production monitoring and control. The term "MRP II" was coined to identify the evolution from *material* resource planning to *manufacturing* resource planning.

Applications targeting departmental became widespread. Aiming at improving the operations in Marketing, Finance, Sales and Payroll, to name a few examples, each of these applications had their own representation and storage of business entities, even for shared concepts such as Customer and Product.

> *Administrative control of a business application along with its required resources gave business managers a degree of freedom and agility. However, by virtue of that distribution, the managers within one line of business could dictate the processes and constraints associated with the development of their own vertical applications to run their own lines of business (Loshin 2009).*

[5] Reorder Point is the level in inventory at which a replenishment order should be placed.

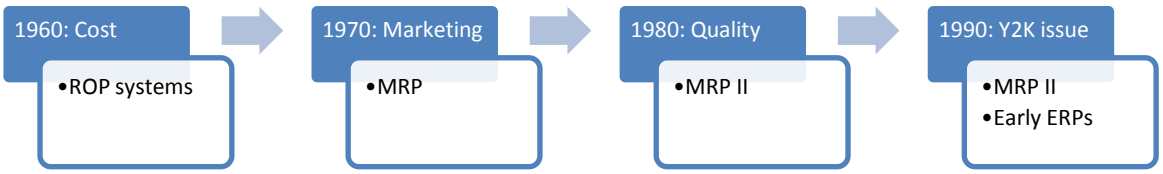

Figure 2 Competitiveness drivers and main systems by decade. Adapted from (Robert Jacobs and Ted' Weston 2007).

In the early 90's, the advent of the CRM concept came with the promise of a single view of customers, which fragmented into several disciplines, largely dictated by vendors (Tuck 2008). Ultimately, there were multiple and inconsistent views of customers in system such as *Sales Force Automation, Analytical CRM* and *Campaign Management Systems.*

In addition, enterprises organized by lines of business (LoB) traditionally managed customer information in a product-centric model, with overlapping customer domains (Dubov 2008). In this scenario, data *ownership,* in addition to systems, is the main driver for redundancies and inconsistencies

Figure 3).

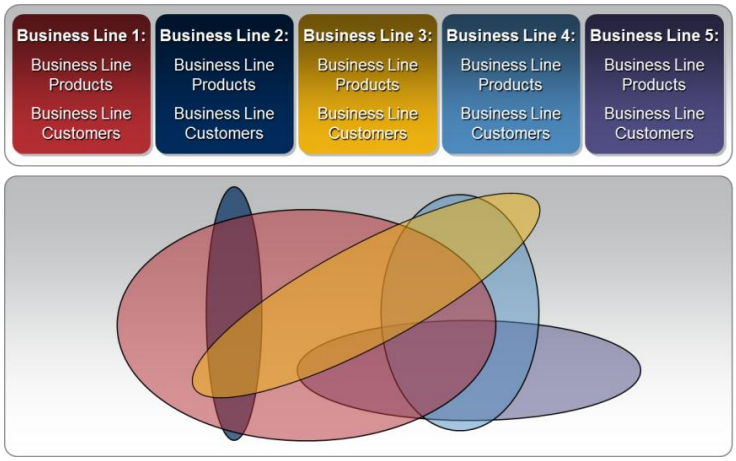

Figure 3 Enterprise Customer Data Managed by Lines of Business and overlapping domains

2.3 Master data management

Master data objects are those "core business objects used in the different applications across the organization, along with their associated metadata, attributes, definitions, roles, connections, and taxonomies" (Loshin 2009). According to (Dreibelbis, Hechler et al. 2007),

MD is the "enterprise-wide data and facts describing key business entities", including facts and relationships among customers, employees, partners, and suppliers, as well as details, facts, hierarchies of products, items, materials, and bill of materials. Master Data Management (MDM) can be defined as an application-independent process which describes, owns and manages core business data entities (Boris and Andreas 2010); a set of policies and procedures for accessing, managing and maintaining a single version of master data and for coordinating the data with subscribing systems across the enterprise for the purposes of maintenance, analysis, and reporting (Ahern 2006).

(Dreibelbis, Hechler et al. 2007) place master data management as one of the three high-level types of data management systems (Figure 4). While transactional management systems support and store data from daily business operations, such as figures and quantities from sales of finished goods, master data systems handle reference data, entities and descriptors shared among business processes.

One classical example of MD is Product data, which in a complete "order to cash" process is consumed in several sub-processes in different systems. From selecting products for placing a purchase order to the invoicing and delivery of goods, specific attributes of the product are used, such as product categories, packaging information, pricing, availability in stock, and so on. In many cases, one single system is unable to fulfil the multiple usages of one master data, leading to key attributes being dispersed in more than one system.

Finally, analytical systems consolidate and present data from both transactional and master data systems, aiming at supporting reporting and analysis. Attributes from master data entities are used to form hierarchies and enable users to see consolidated metrics from different levels of detail, for example across regions, product classification and cost centre structures.

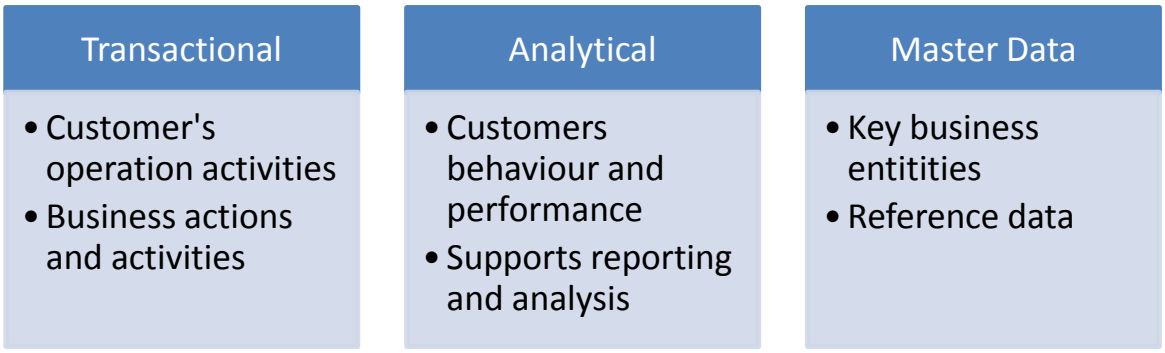

Figure 4 Types of Data Management Systems (Dreibelbis, Hechler et al. 2007)

(Silvola, Jaaskelainen et al. 2011) propose a scheme to define the concept of master data, around the themes of Data (data models, attributes and definitions), Information Systems (for automated sharing and integration) and Processes (data ownership definitions and procedures for cleansing, publishing and protecting data).

Customer information and Product information are the most common forms of master data (Wadehra 2007), and also the ones whose value can be easily measured: customers with wrong or incomplete addresses will not receive their products on time; products with incomplete classification might result in poor inventory management and wrong products being sold. For this reason, Customer Data Integration (CDI) and Product Information Management (PIM) emerged as the first disciplines (accompanied by systems) dedicated to manage a single repository of data which is consumed by processes spanning multiple corporate departments.

The Customer in CDI has a broad spectrum of meanings, which includes a client, contact, party, counterparty, patient, subscriber, supplier, prospect, service provider, citizen, guest, legal entity, trust, and business entity (Berson and Dubov 2007). As disciplines, CDI and PIM are more well-established in comparison to a holistic, process-agnostic, MDM approach and more dedicated literature, for example (Kropsu-Vehkapera, Haapasalo et al. 2009) and (Berson and Dubov 2007).

Account structures or "chart of accounts" (CoA) are also examples of MD. Used for budgeting, planning, forecasting and cost allocation, account hierarchies play in important role in regulatory reporting. Other MD types include Employees, Vendors, Suppliers, Parts, Products, Locations, Contact mechanisms, Profiles, Accounting items, Contracts and Policies (Loshin 2009).

The limited scope of functionally focused or domain centric MDM solutions such as CDI and PIM are attractive because its implementation and benefits are faster to arise (Graham 2010). However, implementing separate solutions will maintain information silos and make it very difficult to manage cross-domain relationships.

MD can serve operational, transactional or analytical purposes. (Dreibelbis, Hechler et al. 2007) propose a taxonomy for *Methods of Use* and architecture patterns. A MDM method of use relates to the ultimate use of master data, namely *Collaborative* (definition, creation and synchronization of master data), *Operational* (consumption of master data by transactional systems) and *Analytical* (consumption of master data as the dimensional source of analytical systems). (Loshin 2009) proposes a similar organization under the name of *Usage Scenarios*, which can be Reference Information Management, Operational Usage and Analytical Usage. Analytic MDM relates to master data items and hierarchies required for aggregation and analysis in business intelligence systems, and the associated requirements for consistent reporting for compliance and corporate performance management. Operational MDM is focused on the distribution or exchange of master data in order to ensure consistency in transactions operations (for example, by means of synchronizing or federating data between different systems).

2.4 Is another system necessary?

Delivering unique, consistent, reliable and traceable data about core business entities is one value proposition of MDM. Historically, these benefits were also intended by ERP, DW and CRM (Moss 2007). Nowadays, ERPs are the main repository of core business entities or *master data*, but why these and other enterprise systems are not appropriate as effective MDM platforms? (Ventana Research 2009; Graham 2010)

> *"I'm sure we've all had the frustration of being told that the person we're talking to can't help us because 'it's not their department' and then getting passed on to someone else to whom we have to explain the whole story again. Or perhaps you've received the sort of marketing mail that I did this week, when my current broadband provider sent me, not one, but two invitations to sign up as a new customer." (Tuck 2008)*

> *"Years ago, a global manufacturing company lost a key distribution plant to a fire. The CEO […]decided to send a letter to key distributors letting them know why their shipments were delayed […].*
> *He wrote the letter and asked his executive team to 'make it happen.' So, they went to their CRM[6], ERP, billing and logistics systems to find a list of customers. The result? Each application returned a different list, and no single system held a true view of the customer." (Fisher 2007)*

The first narrative is a summary of a customer's annoyances caused by improper master data management. The second highlights the problem from a company perspective, where data management inefficiencies will have effects on its financial results.

Both small and large companies face data management issues, starting from misspelled customer's names and incomplete addresses for example. If the operations of a given business are centred on one main system, *fixing* the problem is an easy and straightforward operation. However, the challenge becomes tremendously more complex in global enterprises with different business lines, hundreds of supporting information systems and heterogeneous data management practices inherited from a succession of mergers and acquisitions. Table 2-2 summarizes MDM characteristics and central challenge for organizations of different sizes.

---

[6] Customer Relationship Management. Although a concept, in business jargon it refers to the actual system ("the CRM", "the ERP").

Table 2-2 MDM central challenge for organizations of different sizes

Organization Size	Characteristics	Central Challenge
Small	Small amounts of master data. Data integration is not a top priority.	Creating a plan that will scale with the business.
Mid-size	Data integration starts to become difficult for an organization. Data stewards can be clearly defined.	Implementing effective controls and data stewards.
Large	Huge amounts of master data and system integration. Mostly homogeneous data silos with relatively consistent attributes. Data stewards may now have a full time role.	Building consensus among large number of stakeholders, managing numerous integration points.
Conglomerate	Many disparate businesses that may create many groups of data (i.e., multiple product lines, general ledgers, and so on).	Determining the level at which master data lives.

The following case illustrates the dimensions of the challenge for a global financial services company, which had gone through 90 mergers:

> *"Wachovia, with $706 billion in assets, has completed* 90 mergers *during the past several years. One challenge is reconciling the differences between Wachovia's and the merged companies' data definitions."* (Wailgum 2007)

One common issue with existing applications is the lack of support for custom hierarchies, meaning relationships between business entities for the purposes of data analysis and reporting.

> *[For budgeting, planning and forecasting, we use] Our general ledger's SAP. Our forecasting is done in TM1 (OLAP). Our budgeting is done with arcplan, interfacing into TM1. And we manage all of the master data that we import into the TM1 database with +EDM -- cost centers, profit centers, how they roll up into various hierarchies. (Case in point: Getting the devil out of the details 2006)*

MDM is a trendy technological subject, and experiences a hype which resembles the one from early ERP and Data Warehouses, which was mostly driven by vendors and practitioners who manage to establish a new 'fashion cycle' and leveraging on the momentum generated from a few early adopters (Sammon, Adam et al. 2010). The largest three IT providers, IBM, Microsoft and Oracle, already have products specifically branded with *Master Data* and *Master Data Management*, and the largest service providers already included MDM in their list of expertise (Accenture, HP, Tata Consulting Services).

Against the abundance of tools and systems, authors call companies not to jump the MDM bandwagon without fully grasping the existing data problems in their systems landscape, in order not to repeat the frustrations of rushed ERP, CRM and Data Warehouse initiatives (Tuck 2008; Sammon, Adam et al. 2010), which many companies adopted following an ”everyone is doing it” or “it is time to catch up approach”. Between the progressive view of vendors and the conservative perspective of a few authors, it is worth trying to examine more thoroughly the benefits of MDM initiatives thus understanding the relevance of the research subject.

The historic relevance of MDM relies on attempting to fix a problem that existed for decades, one of data quality and of ensuring a single view of key business entities (White, Newman et al. 2006; Smith and McKeen 2008; Tuck 2008). The issue assumed bigger proportions when companies went global by means of mergers and acquisitions and headquarters has to tackle a new set of data models and established enterprise systems in the new country of operation.

Recent drivers for MDM related initiatives include regulatory compliances such as Sarbanes-Oxley in the United States (2002), Directive on Statutory Audit (2006/43/EC) and, specifically to banks, the Basel II[7] accord in the European Union. Among others, there are regulations for risk management reporting, which demand consolidation of millions of transactions into predefined classes of risk (Kene and Stephen 2007). In a nutshell, the special attention to MDM comes from increasing pressure on regulatory compliance, transparency and accountability (Ventana Research 2009).

MDM assumes additional importance in the light of “cloud” computing model, in which companies have part of its applications and data hosted outside its data center and scaling in an “on-demand” basis. Either by moving existing applications to the cloud or subscribing to new ones, data quality issues tend to arise as a consequence of less control on data models, integration and storage. (Linthicum 2009)

One of the technical challenges in MDM is to consolidate millions of customers from different countries and transactions involving different product lines. A case study with the South African consumer-goods Tiger Brands listed 23 business units using 23 different systems to manage its customer base, a scenario in which data updates could take 160 hours[8]. Another problem with sparse master data is the inability to perform custom roll-ups and drill-downs. The Tiger Brands case, there was a clear demand for hierarchies around customer master data.

---

[7] Basel II imposes standards for risk and capital management, and it relates to MD because of the importance of reference data in segregating exposures into Basel asset classes for regulatory reporting.

[8] According to the case study, the same operations were performed 1 hour after the project completion.

> *"There are a number of data hierarchies around each customer — trade market, channel, region, etc. — but not all are used for every customer in every system. The customer-driven hierarchy allows Tiger to easily evaluate data across these hierarchies, by criteria such as, 'in this region, report on these customers, by trademarketing code', regardless of how the individual store systems track such information."* (Stratature 2005)

Returning to fundamental question proposed at the beginning of this section, the answer partly resides on one key characteristic of functional enterprise systems deployment. In order to be fully tailored, data is modeled according to business processes or functions those systems implement, thus making it less useful for others systems in the organization. (Graham 2010) proposes main five limitations of existing functional systems as MDM applications, as follows:

- Limited master data versioning capabilities
- Inefficient methods of exporting this data into other applications
- Master data is stored in a process-dependent manner
- Inability to properly store hierarchies
- Limited or no ability to model relationships between different data groups

3 THE *HOW* OF MASTER DATA MANAGEMENT

3.1 Promising MDM models

Is despite of a lack of consensus in what constitutes the concept of MDM (Smith and McKeen 2008; Sammon, Adam et al. 2010; Silvola, Jaaskelainen et al. 2011), a common point of view is that MDM is not a technology effort (Fisher 2007), but a technology-enabled business discipline (White, Newman et al. 2006; Hewlett-Packard 2007; Newman and Gall 2009). In other words, technology is an inseparable element of a MDM solution, which has to be introduced as part of a comprehensive plan on data management processes.

Because of the complexities in technology and management challenges related to MDM, a few model or frameworks have been proposed by technology vendors, service providers and practitioners. Published under the names of reference architectures, patterns, best-practices or paradigms, they provide a set of required components or functions that a MDM solution should cover.

(Boris and Kai 2009; Loshin 2009) propose a Functional Reference Architecture, which describes and categorize functions considered necessary for corporate MDM, based on data collected from three rounds of focus groups and interviews with subject matter experts. The model was designed following principles from Business Engineering methodology, which prescribes that business transformations that are based on strategic use of IT should be designed in three different levels (strategic, organizational and system levels), illustrated in .

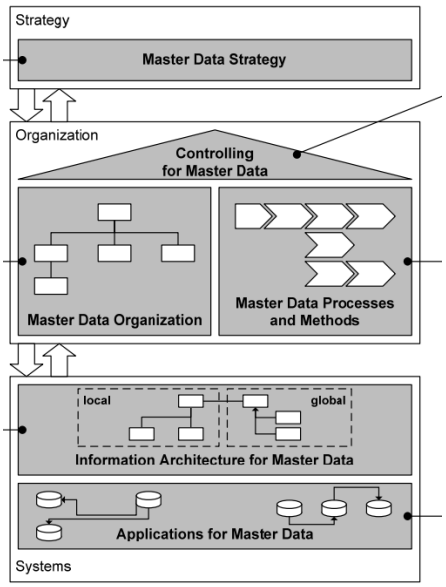

Figure 5 Design areas for corporate master data management (Boris and Kai 2009)

The functions (or *requirements*) are subdivided in three levels, starting with six functional categories at level one. The third and most detailed level comprises 72 discrete functions (Appendix 3 – Functional Reference Architecture for Corporate MDM). This model has four application scenarios, namely *Evaluation*, *Layout Plan*, *Roadmap* and *Information and Experience Exchange*.

The research institute Gartner published a framework called "Seven Building Blocks of MDM", according to which seven main elements should be present in order to guarantee the success of a MDM initiative (Figure 6). Each element or "building block" is accompanied by a list of questions to be answered when planning a MDM solution (**Error! Reference source not found.**).

Building a business case for MDM, in other words, quantifying benefits and acquire commitment from top management, is an essential step and very difficult to execute. MDM is perceived as an infrastructural initiative which lacks business benefits (Dubov 2008; Smith and McKeen 2008).

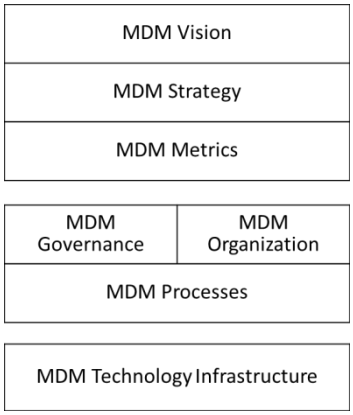

Figure 6 "The Seven building blocks of MDM" (Gartner 2009)

Table 3-1 A comparison between tree MDM models

Model and author	Strengths	Weaknesses
Seven building blocks, Gartner (Radcliffe 2009)	Actionable model. Can be used for early and on-going MDM projects.	Questions are broad. Because of copyright issues, case studies cannot explicitly refer to this model.
Functional Reference Architecture (Boris and Kai 2009)	Actionable model. Verifiable methodology and data sources.	Governance issues are not discussed in depth. Clear focus on the technology side of MDM.
Preconditions for one master data (Silvola, Jaaskelainen et al. 2011)	Simple model, can be used on early stages of MDM.	No guidelines for implementation are provided. Data sample is limited.

3.2 MDM and Governance

The MDM issue is materialized into data repositories. Redundant, incomplete or inaccurate data reside inside databases, which in turn are managed by databases systems executed by computer systems. The tangible, immediate effect of data problems is highly technical; however its origins and solution are broader, and relate to enterprise architecture and processes to manage data. Thus, solving the inconsistences in the actual data is only part of the solution. The most difficult issue is to implement processes that do not allow data issues to be created in the first place. (Radcliffe 2009) highlights that "the greatest challenges will not be technical, but governance-related". "Data governance, as business practice enabled by the right technology, has the potential to create a leap forward in our capability to manage data" and "successful implementations of policy-centric data governance will produce pervasive and long-lasting improvements" (Chen 2010).

(Loshin 2009) emphasizes that MDM is an enterprise initiative, and for this reason "there must be some assurance of stakeholder adherence to the rules that govern participation and information sharing". In addition, master data governance processes should be managed as a subset of a more comprehensive IT and corporate governance (Radcliffe 2009).
Data being managed as an strategic asset is present in businesses with mature data governance models (Chen 2010), in which a "distinct data organization with institutionalized governance processes become a permanent business function".

> *IT governance is defined as the leadership and organizational structures, processes and relational mechanisms that ensure that an organization's IT sustains and extends its strategy and objectives. (Van Grembergen W. 2004).*

The definition focus on tree pillars named *organizational structures*, *processes* and *relational mechanisms*; implementing IT Governance, in a nutshell, is a mixture of those three elements. The author emphasizes that a given IT Governance implementation is contingent on conflicting internal and external factors, thus leading to unique combinations which are applicable only for one company.

According to Loshin (2009), data governance should ensure that data live up to the expectations of all the business purposes, from the perspectives of data stewardship, ownership, compliance, privacy, security, data risks, data sensitivity, metadata management, and MDM. In other words, MDM is an enabler for an enterprise data governance program.

4 RESEARCH DESIGN

4.1 Research method and underlying assumptions

> *All research (whether quantitative or qualitative) is based on some underlying assumptions about what constitutes 'valid' research and which research methods are appropriate. In order to conduct and/or evaluate qualitative research, it is therefore important to know what these (sometimes hidden) assumptions are. (Myers 1997)*

This research was conducted with a qualitative approach, which is suitable for uncovering and understanding a phenomenon about which little is known. As compared to quantitative methods, a limited number of observations is employed, under the justification of 'in-depth studies' or 'thick descriptions' (Ghauri and Gronhaug 2002). In the field of IS research, qualitative methods assume increased importance when its focus shifted from technological to organizational and managerial issues (Myers 1997).

A quantitative method was discarded because the objectives of this research and also because of time and resources constraints. In order to be of any significance, a quantitative method would have demanded a larger data sample. In addition, opting for a quantitative method would have implied narrowing the scope of analysis to one particular aspect of MDM. This research has an explorative orientation, focused on understanding rather than on testing and verifying hypothesis or theories.

Using Järvinen and Järvinen's taxonomy of research methods (Järvinen 2004), this research has simultaneously two approaches, both *conceptual-analytical* (first analyzing basic assumptions about constructs and then identifying theories, models and frameworks used in previous studies) and *innovation evaluation* approach (a realized final state is compared with the desired goal state, eventually by means of specific criteria and measurements).

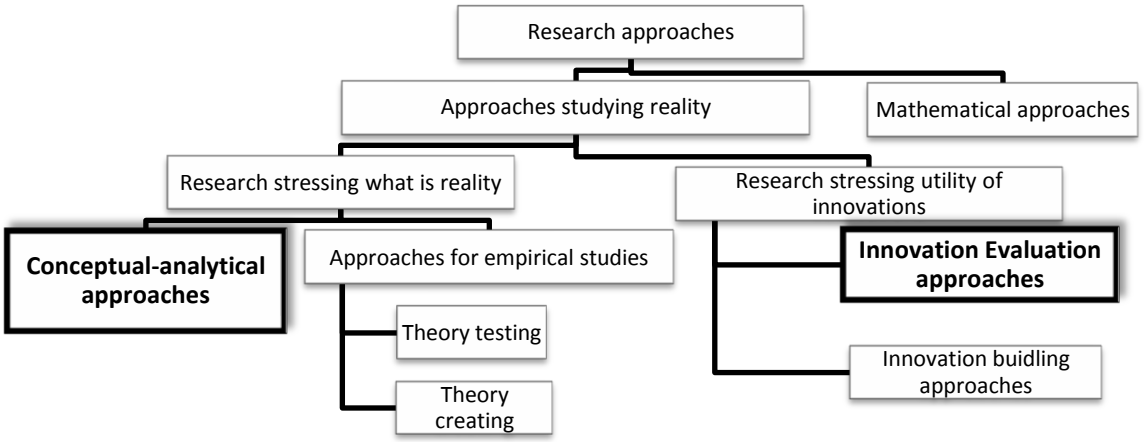

Figure 7 Järvinen and Järvinen taxonomy of research methods (Järvinen 2004), approaches for this research highlighted

In terms of underlying philosophies, this research adopts an interpretive stand, assuming that "access to reality (given or socially constructed) is only through social constructions such as language, consciousness and shared meaning" (Myers 1997). In other words, the researcher is interested at understanding *MDM* phenomena through accessing the meaning participants assign to them[9]. Thus, this research is not premised on *a priori* fixed relationships within phenomena which are normally studied with "structured instrumentation" (Orlikowski and Baroudi 1991). No theories are being tested inferences are drawn from a sample to a stated population, which are characteristics of positivism values in research.[10]

4.2 Research Design

Master Data Management was first studied through existing literature, including both scholar articles and books from practitioners. As previously mentioned, MDM is an emerging

[9] A more detailed discussion can be found in Orlikowski, W. J. and J. J. Baroudi (1991). "Studying Information Technology in Organizations: Research Approaches and Assumptions." Information Systems Research **2**(1): 1-28.

[10] Positivism values used to be the main philosophical ground for IS research in 1980's. A study examining 155 published papers from 1983 to 1988 Ibid. concluded that only 3.2% were considered to have employed an interpretive epistemology.

subject, and limited academic research is available. For this reason, articles from private research institutions such as Gartner and Forrester were also included.

Secondly, two phases of interviews were planned. Phase 1 main goal was to elucidate tacit knowledge in the field of Data Management, in which Ineo has extensive experience accumulated during data migration projects for large Finnish corporations such as Metso, Kone Cranes, Onninen and Wärtsilä. The areas of interest in Phase 1 were:

- Data Management issues, opportunities and benefits
- The relationship between theory and practice in the field of MD
- MDM and Decision Support Systems
- Information Systems for MDM

Phase 2 is aimed at applying the Functional Reference Architecture (Boris and Kai 2009) to Ineo's Master Data project, which at the time of this research was in on-going development. As previously mentioned, the Functional Reference Architecture is organized around six functional categories and the most detailed level comprises 72 functions, which can be mapped to one requirement in an MDM products. A questionnaire was prepared to assess the fit of each functional area using an ordinal scale (Ghauri and Gronhaug 2002), graded in three levels "Fully supported", "Partially supported" and "Not supported". Descriptions for each level were presented to each interviewee at the beginning of the interview, and are listed in Table 4 1. The complete list of questions is included in Appendix 2 – Questionnaire Phase 2. Table 4 1 Description of levels of measurement for Phase 2 questionnaire. Finally, Figure 4.2 illustrates the adopted research process.

Table 4-1 Scale for questions of Phase 2 questionnaire

Level	**Description**
Fully supported	The function is fully present
Partially supported	The function is supported with limitations in scope
Not supported	The function is currently not supported by the software

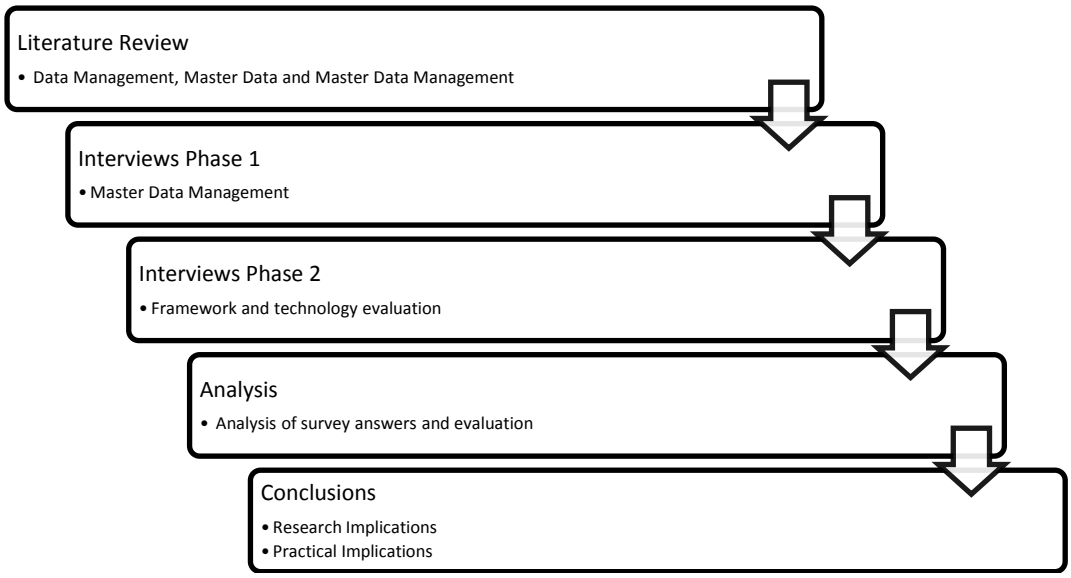

Figure 8 Research process

4.3 Data Collection

This research is mainly based on primary data, which is a favoured option against secondary data because it is more consistent with the particular set of research questions and research objectives of this master thesis (Ghauri and Gronhaug 2002). In addition, the use of primary data is justified by the need of the most current information on subject of MDM.

The data collection technique employed was personal interviews using a semi-structured questionnaire. The method is suitable to this research because it allows elicitation of tacit knowledge and in-depth observation of a phenomenon. Interviews were recorded and performed with one informant at time. The time reserved for the interviews was 60 minutes each.

4.4 Informants of the study

The interviews were conducted with four managers from Ineo (named A to C), one system analyst from Metso Minerals Brazil (D) and one manager from Wärtsilä Finland[11] (E).

[11] Wärtsilä is global company headquartered in Finland whose main line of business are Ship Power, Services and Power Plants (http://www.wartsila.com/en)

Informants from Ineo have from 5 to 10 years of experience in data management projects and currently hold senior management positions.

Informant A is currently Chief Executive Officer, and has been a project manager in data migration projects for ERP implementations and ERP consolidation. Informant B has a similar experience, and its currently Chief Project Officer. Informant C has 10 years in leading BI and DW project, is the Concept Owner for Business Intelligence at Ineo, and also the CFO in charge. Informant D worked as master data coordinator for product data in SAP implementations in Mexico, Chile, Peru and Brazil, more specifically in Product Data Management (PDM) module. In despite of informant D not providing answers to all the questions, his contribution was not discarded as a whole, because the answered questions were judged as valuable to this research. Topics without answer are listed as 'Not answered' in research results.

Finally, Informant E is currently a Senior Information Manager, who has worked in the Business Information Centre during three years. His responsibilities included MDM process development and the specification of KPI's to measure data issues and link them to possible financial indicators. A summary of informants' profile is presented in Table 4-2.

Table 4-2 Informants for Phase 1, Master Data Management

Informant	Experience	Role
A, Ineo Oy	Data harmonization and data migration projects	Project Manager, currently CEO
B, Ineo Oy	Data harmonization and data migration projects	Project Manager, currently CPO
C, Ineo Oy	Business intelligence	Architectural Consultant, CFO
D, Metso	SAP implementation in PDM modules	Master data coordinator
E, Wärtsila	MDM processes development	Senior Information Manager

4.5 Analysis

Primary data from the interviews was analysed using a qualitative approach. Notes taken during the interviews were reviewed after listening to the recorded audio, and the key topics were analysed and summarized into four categories, as follows:

- Data Management
- Master Data and Master Data Management
- Processes and Governance

- MDM and Information Systems

The aforementioned categories correlate respectively to groups B to E in questionnaire one (listed in Appendix 1 – Questionnaire Phase 1).

For phase 2 of the research, a more quantitative approach was adopted. Answers were counted and categorized according to functional areas as proposed by (Boris and Kai 2009).

4.6 Research limitations

The research was conducted with a limited set of employees from two companies. The conclusions of this research require further investigation in order to assess its applicability in different contexts. The practical outcomes of this research were tailored to Ineo Oy and its Research and Development project. The research and thesis writing was conducted in five months, which imposes constraints on an exhaustive literature review and a data collection within a wider universe. Possible areas for further exploration are discussed in chapter 0, The underlying data issues in Master Data Management discipline are complex and largely underestimated. As it was stated by one informant, one given set of data might require several cycles of cleansing and enrichments, due to organizational changes that current systems simply can cope with. For example, if one physical warehouse is deactivated, making the required changes in enterprise systems is a long and complex undertaking, due to relationships between material classifications in global and local ERP instances, the associate vendors and existing production planning records.

In order to fulfil the business and governance requirements proposed by current literature, the technological element in a master data management solution has to implement several layers of intricate integration and synchronization mechanisms. In this sense, one challenge is a data management mechanism which is domain-independent, and which supports complex relationships and hierarchy management. For example, the current Customer Data Integration and Product Information Management solutions are tailored for customer and product, and are built with assumptions on how customer and product data is produced and consumed. A holistic MDM solution has to accommodate *any* master data entity.

Companies considering embarking the MDM trend will find technological options from most of their existing providers for other corporate systems. The established vendors of ERP and CRM are favoured by corporate strategies looking for economy of scale in IT sourcing, in which one vendor is the preferred provider for corporate systems. However, it is expected that the previous experiences with ERP and DW implementations, combined with increased scrutiny on IT investments, will make companies more cautious of pure technology-based solutions. The main message from all the references found in this research is clear: do not approach MDM with technology.

Research Implications.

The two research questions in the process and data governance received vague and broad answers, which hindered further analysis on MDM and data governance. The two proposed reasons for this are the inherent complexity in the subject and a possible limited experience on the subject. For the former, formulating the questions in a more specific way could have led to better answers. For the second reason, a pre-assessment of informants could have depicted in which research topics they had experience.

5 RESEARCH RESULTS

5.1 Data Management

Data management projects are usually initiated as a part of larger enterprise systems implementation or to accommodate organizational changes such as mergers and acquisitions and reallocation of accounts structures (informants A and B, Ineo). For example, when companies acquire others, there is need to merge ERP instances to the local business. In addition, global enterprise architecture changes related to unification of business processes might also trigger data management projects, aiming at harmonize and unify data from different systems (informant B). Informant C emphasized that ERP and other "line of business" applications are "point in time" systems, which do not provide history tracking of data records.

The challenges in implementing successful data management projects are related to a lack of perception from management on the real effort required to solve data problems and the amount of resources allocated (informant A, Ineo). It is a time consuming project, and understanding the data models is one of the most difficult steps, both in terms of logical models and actual implementation of data harmonization. The problem is not understood correctly by companies, which over-emphasize de-duplication of business entities and ignore the need for harmonization of attributes, which is a requirement for reporting and analytics. Finally, core transaction processes, such as cash to order, receive most of the attention, and data management is faced in a reactive manner, without specific data management processes (informant C, Ineo).

In terms of costs and benefits of data management projects, it was a common perception that it is difficult to calculate the costs of poor quality data. In addition, data management projects are usually part of a more complex enterprise project associated to a system implementation, thus costs and benefits being calculate for a technology implementation rather than data quality. It was mentioned that it is easier to calculate the additional costs incurred for particular cases, for example when duplicate materials lead to excesses in inventory or inconsistent product data led to a wrong products being delivered to a customer (informant E, Wärtsilä). However, from top management perspective, these particular cases are not enough to justify the investments required to permanently solve the problem.

Table 5-1 Data Management drivers, challenges and costs-benefit analysis

Informant	Main Drivers	Challenges	Cost-benefit analysis
A	Organizational changes	Data is not the priority. Effort and resources required are overlooked.	It is not common to quantify costs and benefits for data
B	Rationalization of resources, merging of ERP instances, global architecture changes	Understand data models, implement data harmonization on keys and attributes	Measured on ERP and BI implementations, but do not consider data
C	Consolidation roll-up of customer and product structures; change over time of hierarchical structures	No processes are defined for data management	Benefits are indirect; MDM is the pre-condition for the realization of the ERP business case
D	Data quality, modifications in legacy systems, new management procedures	Proper allocation of resources, definition of data ownership, continuous data-quality verification	Extremely hard to measure. Can include: efforts in data maintenance, savings with fines, costs with data migration
E	Poor customer and material information	Solutions are too specific	Difficult to measure. Easy for particular instances

5.2 Master data and master data management

From the interviews it was possible to conclude that even though MDM is an evolving concept, there is already a shared understanding that is a complex business problem, which demands a solution rooted in process and governance, and not in technology. It was commented by all the informants that literature on the subject are over simplistic and cover only the broad, generic cases.

When being asked about information assets, one informant stated that in most of the cases, data is treated as being part of IT systems and teams.

Informant	A	B	C	D	E
Theory and practice	Theory is accurate	Theory is simple; reality is more complex	Theory is accurate for “Global” entities	No answer provided	Concepts in theory are correct, however only “easy cases” are being discussed
Maturity level	Practices need to evolve	Concept is fragmented, and involves different disciplines	Currently oversimplified, process level requirements are not understood	Practices do exist, but coordination and organizational issues are the biggest challenges	Low; team is mostly “fixing problems”; no executive ownership
Information assets	Executive levels realize the value, practices are lagging behing	Global, shared data	Data that is meaningful regardless of context in which it was created and the person who entered	No answer provided	Data is still considered part of IT
Development	Towards efficiency; best practices will be defined	Will evolve from an ”abstract nuisance” to an established corporate discipline	Market consolidation and best practices established	Awareness on the subject will arise; tools will provide better data stewardship and integration	Not a buzzword anymore; practices and IT

5.3 Process and data governance

It was a common opinion that establishing processes is a pre-requisite for successful data management initiatives. In relation to corporate governance, assigning responsibilities for business entities’ data (data ownership) was considered difficult. Informant C argued that MDM vendors do not discuss the challenges in data management because of a lack of knowledge in business process and the commercial interests.

Informant D emphasized that coordination with corporate governance in MDM project assumes varying degrees of importance. There are situations in which only data is affected, but when changes happen in the way people do their daily work, then establishing and communicating governance principles is essential for the project to succeed.

Informant	Emphasis on processes	Relation to IT and Corporate Governance
A	Mature processes are present only in global MDM strategies	MDM is established with “pilots”
B	No answer provided	No answer provided
C	Crucial, but overlooked by vendors	MDM follows current governance trends towards centralized data
D	Processes are usually discussed but not deeply	Varying degrees of coordination to corporate governance;
E	Process and governance must come first	Difficult to define “who owns the data”

5.4 Information Systems for Data Management and MDM

Information systems for master data management in large corporations are often times applications developed “in-house”, tailored for point-to-point integration between legacy systems. This is a consequence of the limitations in current enterprise systems in handling end-to-end data migration (Informants A and B, Ineo). Informant E stated that the current technology tackles only specific issues and do not provide a comprehensive solution, able to cover all the aspects of MD lifecycle. He also suggests that there a clear business opportunity for *master data providers* – companies selling subscriptions to a shared, accurate and updated central repository of Customer master data for example.

Another limitation is the lack of custom hierarchies among business entities hierarchies, which is a requirement for advanced analysis scenarios. One example is the hierarchy “Customer, Installation, Engine, Part, Vendor”, which currently cannot be maintained in the ERP. (Informant E, Wärtsilä).

Informant C stated that ERPs and other LoB applications have inbuilt data structures that do not accommodate other usages of the data it stores. In other words, they cannot describe and store a business entity in a way which fulfils the needs of other enterprise systems.

The main key points for this section are summarized in Table 5-2.

Table 5-2 Information Systems for Data Management: current issues and expected evolution

Informant	Existing DM technology	Evolution
A	Current tools are too complex and expensive. In complex landscapes, several small solutions were developed to fulfill local needs of legacy systems.	Better alignment of tools and business processes.
B	Existing MDM packages are tailored for systems of the same vendor. Workflow functionality does not cover all existing systems.	Inclusions of Data Migration and better workflow functions. Tools will requires less IT participation and will be “business-friendly”
C	ERPs and other LoB applications cannot describe and store a business entity for general use	Market consolidation after acquisitions and success implementations
D	Not answered	Not answered
E	Targeted at very limited scope Limited support for custom hierarchies Constrained by SAP models	Master data will move away to different systems

5.5 CIAM Suite assessment

This chapter covers the results of Phase 2 of this research. The Functional Reference Architecture (Boris and Kai 2009) was enhanced in order to be applied as a questionnaire. The adopted approach was to convert each function (or requirement) into a scale of “Fully Supported”, “Not Supported” and “Supported”. The informant was asked to fill in the questionnaire by simply marking with a cross the most appropriate level of compliance to each requirement. For example, if the informant believed that “Conditional Entries” was a feature offered by the product under development, he would choose “Fully Supported”, and so on. Figure 9 shows the first functional category and functional requirements with corresponding answers. The full questionnaire with answers is presented in Appendix 4.

Functional Categories and Functions	Fully Supported	Partially Supported	Not Supported
Master Data Lifecycle Management	7	2	0
Data Creation	3	0	0
Conditional Entries	X		
Bulk Editing	X		
Plausibility Check	X		

Figure 9 A sample of questionnaire from phase 2

The results were counted and summarized into the six functional categories proposed by the model (Figure 10). According to the assessment, the *CIAM Suite* product fully supports 43% of the requirements (last bar in the chart). The functional area with least support is *Metadata Management and Master Data Modelling*, which has 33% of features classified as not supported. *Master Data Integration* is the area with the highest percentage of "Partially Supported" features (9 out of 14 requirements).

Figure 10 CIAM compliance to Functional Architecture Reference, by Functional Category

Regarding the methodology for research phase 2, the informant expressed a positive attitude towards the grouping of feature, and emphasized that employing an assessment tool that is vendor- and process-independent brings insights into the current development processes, which might be biased by the experience of those responsible for its high-level definitions. In addition, the informant highlighted that a few functional areas descriptions are vague, and required further reading of the original reference, in which they are described in detail. As a benchmark tool, the Functional Reference faces the challenge of covering

functionality that might have different names and descriptions according to vendors branding and focus on one aspect or another of MDM.

6 CONCLUSIONS

The underlying data issues in Master Data Management discipline are complex and largely underestimated. As it was stated by one informant, one given set of data might require several cycles of cleansing and enrichments, due to organizational changes that current systems simply can cope with. For example, if one physical warehouse is deactivated, making the required changes in enterprise systems is a long and complex undertaking, due to relationships between material classifications in global and local ERP instances, the associate vendors and existing production planning records.

In order to fulfil the business and governance requirements proposed by current literature, the technological element in a master data management solution has to implement several layers of intricate integration and synchronization mechanisms. In this sense, one challenge is a data management mechanism which is domain-independent, and which supports complex relationships and hierarchy management. For example, the current Customer Data Integration and Product Information Management solutions are tailored for customer and product, and are built with assumptions on how customer and product data is produced and consumed. A holistic MDM solution has to accommodate *any* master data entity.

Companies considering embarking the MDM trend will find technological options from most of their existing providers for other corporate systems. The established vendors of ERP and CRM are favoured by corporate strategies looking for economy of scale in IT sourcing, in which one vendor is the preferred provider for corporate systems. However, it is expected that the previous experiences with ERP and DW implementations, combined with increased scrutiny on IT investments, will make companies more cautious of pure technology-based solutions. The main message from all the references found in this research is clear: do not approach MDM with technology.

6.1 Research Implications

As many authors emphasize, there is little research on the field of MDM (Smith and McKeen 2008; Silvola, Jaaskelainen et al. 2011), and this master's thesis initial ambition was to consolidate the most current information about the subject from both academic and practitioners perspectives.

The problems that MDM promises to solve are undisputable. Organizational structures and business processes have to continuously evolve to accommodate the strategic decisions on business operations. Driven by competitive forces or the basic principle of increasing profits, this evolution will always push for changes in enterprise systems or its underlying data. In order to circumvent limitations in the existing systems design, data about key business

entities is manipulated without defined rules and without considering long-term effects in the data usage chain, leading to uncountable duplicates and inconsistencies in attributes usage.

One area lacking further research is how to build the business case for MDM, in other words, how to gather metrics about direct and indirect costs associated to poor master data quality and about all potential cost savings. In one extreme, vendors might state that "selling the business case is as easy as 1-2-3" (Wadehra 2007), and on the other end, managers say "the business case for MDM is *never* compelling enough" (Smith and McKeen 2008). In addition, the interviews conducted during the course of this research also reinforced that financial benefits of MDM projects are complex to measure. One practical example was one occasion in which products were sold by an incorrect price. Even though the additional costs associated were easy to commensurate, the investments required to permanently fix the problem by far outstand the benefits in a foreseeable future. This analysis could be performed within one subsidiary of a global enterprise by means of a qualitative case study on the business processes combined with quantitative methods to measure the aforementioned issues.

6.2 Practical implications

The most significant practical implication for Ineo Oy is a framework which can be used to evaluate the CIAM product. The Functional Reference Architecture for Corporate Master Data Management (Boris and Kai 2009) is an application-agnostic framework, derived from focus-group studies with senior executives from German companies. The application of the Functional Reference framework to Ineo solution can refine its scope delimitation, by means of insights from end-users from a variety of industries on what is expected from MDM solutions. In addition, it can serve as a tool to strengthen Ineo marketing strategy as being fully compliant to one public, application-independent evaluation model.

Another common shared view among interviewees and references was the importance of a scalable approach to MDM projects. MDM capabilities should be built up in phases, with clear added-value in each phase (Radcliffe 2009). "Identify small, quick-wins", focusing on one type of data (Smith and McKeen 2008); "but start small and deliver incremental benefits quickly" (Tuck 2008). Thus, given the complexity of Ineo's MDM solution, the company should strongly consider how customers could acquire and deploy smaller pieces of the solution, in order to create a sustainable implementation roadmap.

From the interviews and from the researched literature it was possible to conclude that even though MDM is an evolving concept, there is already a shared understanding that is a business problem that cannot be tackled only with technology. Software applications will only help when appropriate processes are in place. Thus, Ineo should consider how a

potential customer can assess its “readiness” for CIAM Suite, for instance, by providing assessment tools or consulting services.

One highlight from the interview with Manager E was the emphasis on the business opportunity for data providers which would concentrate and organize master data for customers and products, to which other businesses would subscribe and incorporate in their processes.

7 SUMMARY

After decades of mergers and acquisitions and successive IS trends such as CRM, ERP and DW, the data in enterprises systems is scattered and filled with inconsistencies. Global organizations face the challenge of addressing local uses of shared business entities such as customer and product, and at the same time have a timely, consistent, unique, and consolidate view of financial indicators. In addition, current enterprise systems do not accommodate the pace of organizational changes and immense efforts are required to maintain data. When it comes to systems integration, ERPs are considered "closed" and expensive. Data structures are complex and the "out-of-the-box" integration options which are offered by these systems are not based in industry standards, thus expensive and time-consuming projects are undertaken.

Master Data Management emerges as one discipline focused on ensuring long-term data consistency. Presented as a technology-enabled business discipline, it emphasizes business process and governance to model and maintain the data related to key business entities. There are immense technical and organizational challenges to accomplish the "single version of the truth" MDM mantra. Adding one central repository of master data might prove unfeasible in a few scenarios, thus an incremental approach is recommend, starting from areas most critically affected by data issues.

This research aimed at understanding the current literature on MDM and contrasting it with views from practitioners. The data collected from interviews revealed details on the complexities of data structures and data management practices in global organizations, reinforcing the call for more in-depth research on organizational aspects of MDM. Most of what has been written so far can be applied to a broad, generic or "global" levels of master data, in other words, the basic set of attributes which are meaningful for any data application across the globe, such as product or customer classification. The most difficult piece of master data to manage is the "local" part, the attributes related to the sourcing and storing of materials in one particular warehouse in The Netherlands or a complex set of pricing rules for a subsidiary of a customer in Brazil.

8 REFERENCES

Ahern, I. (2006). High-Octane MDM: How Enterprise Dimension Management Boosts BPM's Performance. Business Performance Management Magazine, Penton**:** 32-35.

Berson, A. and L. Dubov (2007). Master Data Management and Customer Data Integration for a Global Enterprise. Emeryville, CA, USA, McGraw-Hill Professional Publishing.

Boris, O. and R. Andreas (2010). Organizing master data management: findings from an expert survey. Proceedings of the 2010 ACM Symposium on Applied Computing. Sierre, Switzerland, ACM.

Boris, O. and H. Kai (2009). Functional Reference Architecture for Corporate Master Data Management, Institute of Information Management, University of St. Gallen.

Case in point: Getting the devil out of the details (2006). Business Performance Management, Penton Media, Inc.

Chen, W. (2010, 19.03.2011). "Kalido Data Governance Maturity Level." from http://www.kalido.com/Collateral/Documents/English-US/Kalido-Data-Governance-Maturity-Model-1-5.pdf.

Cooper, H. (1988). "Organizing knowledge syntheses: A taxonomy of literature reviews." Knowledge, Technology & Policy **1**(1): 104-126.

Dreibelbis, A., E. Hechler, et al. (2007) "Information service patterns, Part 4: Master Data Management architecture patterns." **Volume**, DOI:

Dubov, L. (2008). Implementing Enterprise Information Programs with MDM-CDI & SOA, Data Administration Management Association of New York, DAMA-NY.

Fink, A. (2005). Conducting research literature reviews : from the Internet to paper. Thousand Oaks, CA :, Sage Publications.

Fisher, T. (2007). "Demystifying Master Data Management." Retrieved 23.04.2011, 2011, from http://www.cio.com/article/106811/Demystifying_Master_Data_Management.

Ghauri, P. and K. Gronhaug (2002). Research methods in business studies : a practical guide. Harlow :, Financial Times Prentice Hall.

Graham, T. (2010). "Master Data Management from a Business Perspective." Retrieved 24.04.2011, 2011, from http://download.microsoft.com/download/D/B/D/DBDE7972-1EB9-470A-BA18-58849DB3EB3B/MasterDataManagementfromaBusinessPerspective.pdf.

Greenemeier, L. and P. McDougall. (2006, February 6, 2006). "Credibility Of Analysts. Do tech vendors wield influence over IT research? You bet--but how much of it is a matter of perspective?" Global CIO Retrieved March 8, 2011, from http://www.informationweek.com/news/global-cio/showArticle.jhtml?articleID=178601879#.

Hewlett-Packard (2007). The seven deadly sins of master data management, Hewlett-Packward**:** 7.

Järvinen, P. (2004). On Research Methods. Tampere, Tamperen Yliopisto Oy.

Kene, R. and A. Stephen. (2007). "Crossing the Reference Data Hurdle in Basel II Implementation." Information Management Special Reports Retrieved 18.04.2011, 2011, from http://www.information-management.com/specialreports/20070918/1091348-1.html?pg=1.

Kropsu-Vehkapera, H., H. Haapasalo, et al. (2009). "Product data management practices in high-tech companies." Industrial Management & Data Systems **109**(6): 758-774.

Linthicum, D. (2009). "MDM Becoming more critical in light of cloud computing." Retrieved 28.04.2011, from http://www.ebizq.net/blogs/linthicum/2009/09/mdm_becoming_more_critical_in.php.

Loshin, D. (2009). Master Data Management. Oxford, Elsevier Inc.

Moss, L. (2007). "Critical Success Factors for Master Data Management." Cutter IT Journal **20**(9): 7-12.

Myers, M. D. (1997). "Qualitative Research in Information Systems." MIS Quarterly **21**(2): 241-242. MISQ Discovery, archival version, June 1997, http://www.misq.org/discovery/MISQD_isworld/. MISQ Discovery, updated version, last modified: February 17, 2011 www.qual.auckland.ac.nz.

Newman, D. and N. Gall (2009). A Master Vocabulary for Ensuring Consistency Across Information Management Disciplines Gartner Research.

Orlikowski, W. J. and J. J. Baroudi (1991). "Studying Information Technology in Organizations: Research Approaches and Assumptions." Information Systems Research **2**(1): 1-28.

Radcliffe, J. (2009) "The seven building blocks of MDM: A framework for success." **Volume**, DOI:

Randolph, J. (2009). "A Guide to Writing the Dissertation Literature Review." Practical Assessment, Research & Evaluation **14**(13).

Robert Jacobs, F. and J. F. C. Ted' Weston (2007). "Enterprise resource planning (ERP)--A brief history." Journal of Operations Management **25**(2): 357-363.

Sammon, D., F. Adam, et al. (2010). Making Sense of the Master Data Management (MDM) Concept: Old Wine in New Bottles or New Wine in Old Bottles? Proceeding of the 2010 conference on Bridging the Socio-technical Gap in Decision Support Systems: Challenges for the Next Decade, IOS Press.

Silvola, R., O. Jaaskelainen, et al. (2011). "Managing one master data – challenges and preconditions." Industrial Management & Data Systems **111**(1): 146-162.

Smith, H. A. and J. D. McKeen (2008). "Developments in Practice XXX: Master Data Management: Salvation Or Snake Oil?" Communications of AIS **2008**(23): 63-72.

Stratature. (2005). "With +EDM, Tiger Brands Tames the Customer Data Jungle." from http://www.information-management.com/issues/20050901/1035589-1.html.

Tuck, S. (2008). Is MDM the route to the Holy Grail? Journal of Database Marketing & Customer Strategy Management**:** 218-220.

Van Grembergen W., D. H. S. (2004). "IT Governance and its mechanisms." Information Systems Control Journal **1**.

Ventana Research. (2009). "New Strategies for Managing Master Data." Retrieved 25.04.2011, from http://www.ventanaresearch.com/research/category_new.aspx?id=1314.

Wadehra, A. (2007). "The ABCs of Master Data Management: Architecture, Business Case, and Customer." Business Intelligence Journal **12**(1): 33-40.

Wailgum, T. (2007). "Master Data Management: Truth Behind the Hype." Retrieved 23.04.20111, from http://www.cio.com/article/110558/Master_Data_Management_Truth_Behind_the_Hype.

White, A., D. Newman, et al. (2006). Mastering Master Data Management, Gartner Research.

APPENDIX 1 – QUESTIONNAIRE PHASE 1

Part A. Interviewee information

1) Interviewee job title.
2) Description of your experience with Data Management (duration, types of projects and industries).

Part B. General questions on Data Management

1) In your opinion, what are the problems or opportunities that lead companies to initiate Data Management projects?
2) What are main challenges in implementing successful Data Management projects? Please consider technology and human factors.
3) How do you describe the process to measure the costs and benefits related to data management issues?

Part C. General questions on Master Data (MD) and Master Data Management (MDM).

1) In your understanding, how MD/MDM theory and practice are related?
2) What is your perception on the current conceptual maturity level of Master Data Management, in terms of definitions and practices?
3) Which types of information can be considered "information assets"?
4) How do you think these terms and related concepts and practices will develop in the next 5 years?

Part D. Processes and Data Governance

The literature places a significant importance on processes and governance for MDM. This section aims at understanding how these two aspects are being addressed.

1) Considering the projects in which you have been involved, what is the emphasis on processes? Do you have examples that can be considered successful and unsuccessful cases?
2) According to your experience, how Data Management and Master Data initiatives or projects relate to IT Governance and or Corporate Governance?

Part E. Information Systems

This section relates to your experience with technology ("software", "solutions", and "tools").

1) According to your experience, to which extent the current technology fulfils existing Data Management issues? Which needs are not currently addressed by technology?
2) How do you think this technology will develop in the next 5 years?
3) In your organization, how Master Data/Data Management activities relate to Decision Support Systems? (organizational structures, systems integration etc.)

APPENDIX 2 – QUESTIONNAIRE PHASE 2

Functional Categories and Functions	Fully Supported	Partially Supported	Not Supported
Master Data Lifecycle Management	7	2	0
Data Creation	3	0	0
Conditional Entries	X		
Bulk Editing	X		
Plausibility Check	X		
Data Maintenance	2	1	0
Check-out	X		
Bulk Editing	X		
Plausibility Check		x	
Data Deactivation	1	0	0
Bulk Editing	x		
Data Archiving	1	1	0
Archiving	x		
History Control		x	
Metadata Management and Master Data Modelling	7	3	6
Data Modelling	2	1	2
Data Model Editing	x		
Graphical Modelling			x
Classification			x
Support of Business Standards		x	
Data Model Version Control	x		
Model Analysis	0	1	3
Dependency Analysis			x
Data Type Recognition		x	
Primary and Secondary Key Recognition			x
Relationship Recognition			x
Metadata Management	5	1	1
Business Rules Documentation	x		
Glossary/Dictionary	x		
Metadata Import	x		
Mandatory Fields Administration	x		
Metadata Publication		x	
Metadata Transport	x		
Metadata Visualization			x

Functional Categories and Functions	Fully Supported	Partially Supported	Not Supported
Master Data Quality Management	3	8	3
Data Analysis	1	1	2
Compliance Verification	x		
Graphical Analysis		x	
Plausibility Lists			x
Profiling			x
Data Enrichment	1	4	0
External Reference Data	x		
Classification Schemes		x	
Measuring Units		x	
Multilingual Capability		x	
Management of Unstructured Data		x	
Data Cleansing	1	3	1
Delta Import	x		
Duplicate Recognition		x	
Pattern Recognition		x	
Plausibility Check		x	
Spelling Check			x
Master Data Integration	5	9	0
Data Import	1	3	0
Delta Import		X	
Import Formats	x		
Connectors		x	
Virtual Integration		x	
Data Transformation	1	0	0
Field Split	x		
Field Merge		x	
Data Type Conversion		x	
Pivot Tables		x	
Data Export	3	0	0
Search Based Data Selection		x	
Delta Export		x	
Export Formats		x	
Connectors	x		
Limitation	x		
Preview	x		

Functional Categories and Functions	Fully Supported	Partially Supported	Not Supported
Cross Functions	10	3	2
Automation	1	3	1
Automated Enrichment	x		
Automated Export		x	
Automated Import		x	
Cross-Function Automation		x	
Push and Pull Mechanisms			x
Reports	4	0	0
Data Quality Reports	x		
Usage Statistics	x		
Job Monitoring	x		
Audit Support	x		
Search	3	0	0
Dynamic Value Search	x		
Free Search	x		
Fuzzy Search	x		
Workflow Management	2	0	1
Bundling of Activities			x
Graphical Workflow Modelling	x		
Create/Maintain Workflows	x		
Administration	2	1	1
Data History Management	1	1	0
Data Lineage	x		
Last User		x	
User Management	1	0	1
User Interface Design			x
Roles and Rights	x		

APPENDIX 3 – FUNCTIONAL REFERENCE ARCHITECTURE FOR CORPORATE MDM

A	1	2	3	4
Master Data Lifecycle Management	**Data Creation**	**Data Maintenance**	**Data Deactivation**	**Data Archiving**
	Conditional Entries	Check-out	Bulk Editing	Archiving
	Bulk Editing	Bulk Editing		History Control
	Plausibility Check	Plausibility Check		

B	1	2	3
Metadata Management and Master Data Modeling	**Data Modeling**	**Model Analysis**	**Metadata Management**
	Data Model Editing	Dependency Analysis	Business Rules Documentation
	Graphical Modeling	Data Type Recognition	Glossary/Dictionary
	Classification	Primary and Secondary Key Recognition	Metadata Import
	Support of Business Standards	Relationship Recognition	Mandatory Fields Administration
	Data Model Version Control		Metadata Publication
			Metadata Transport
			Metadata Visualization

C	1	2	3
Master Data Quality Management	**Data Analysis**	**Data Enrichment**	**Data Cleansing**
	Compliance Verification	External Reference Data	Delta Import
	Graphical Analysis	Classification Schemes	Duplicate Recognition
	Plausibility Lists	Measuring Units	Pattern Recognition
	Profiling	Multilingual Capability	Plausibility Check
		Management of Unstructured Data	Spelling Check

D	1	2	3
Master Data Integration	**Data Import**	**Data Transformation**	**Data Export**
	Delta Import	Field Split	Search Based Data Selection
	Import Formats	Field Merge	Delta Export
	Connectors	Data Type Conversion	Export Formats
	Virtual Integration	Pivot Tables	Connectors
			Limitation
			Preview

E	1	2	3	4
Cross Functions	**Automation**	**Reports**	**Search**	**Workflow Management**
	Automated Enrichment	Data Quality Reports	Dynamic Value Search	Bundling of Activities
	Automated Export	Usage Statistics	Free Search	Graphical Workflow Modeling
	Automated Import	Job Monitoring	Fuzzy Search	Create/Maintain Workflows
	Cross-Function Automation	Audit Support		
	Push and Pull Mechanisms			

F	1	2
Administration	**Data History Management**	**User Management**
	Data Lineage	User Interface Design
	Last User	Roles and Rights

APPENDIX 4 – DESCRIPTION OF GARTNER FRAMEWORK FOR MDM

MDM Element	Description	Questions to ask
Vision	Needs to reflect business vision; describes "what" MDM looks like and "why" it needs to be created	What business vision needs MDM as an enabler? Who owns that business vision? What does the supporting MDM vision look like? Who owns the MDM vision?
Strategy	"How" to achieve the vision; goals, scope, life-cycle	Which domains will be addressed? What's the current MDM maturity and what's the to be state?
Metrics	Identify KPIs that relate to business goals.	What are the measurable objectives for your MDM? How are you communicating them? How will you know how success looks like? Have you made the connection to business value?
Governance	Creation of a governance framework that is part of corporate and IT governance	How will create and maintain an MDM governance organization? How will the governance structure enforce its will?
Organization	Identify needs and responsibilities of MD users; planning of communication, change management and training	Who creates and consumes master data? What are their roles, and how the program will change their work?
Processes	Deciding on "must have" processes	What MDM processes you will need? What business process will the master data life-cycle support?
Technology Infrastructure	Decide which technologies are essential for the MDM vision and strategy; decide on "build or buy"	What IT infra-structure do you need? How does it fit into the overall information architecture? What architectural and implementation styles are appropriate?