Master Data Management: Its importance and reasons for failed implementations

Panagiotis Lepeniotis

A thesis submitted in partial fulfilment of the requirements of

Sheffield Hallam University

for the degree of Doctor of Philosophy

Volume 1

January 2020

Candidate Declaration

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.

2. None of the material contained in the thesis has been used in any other submission for an academic award.

3. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.

4. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.

5. The word count of the thesis is 73,600

Name	*Panagiotis Lepeniotis*
Date	*January 2020*
Award	*PhD*
Faculty	*SHU Faculty of Arts, Computing, Engineering and Sciences*
Director(s) of Studies	*Dr Simon Polovina & Dr Sohrab Moshiri*

Abstract

This research aims to identify the useful impact of Master Data Management (MDM) on a Business Transformation Programme (BTP).

A BTP consists of three distinct phases. The first phase is the selection of the appropriate set of application systems as well as the introduction of new business processes across multiple lines of business and different channels. The second phase is the implementation of the new application systems and the data migration process. The third and final phase is the transition from the legacy application systems and business processes to the newly defined framework of processes and technologies that ensure the business and data continuity. MDM encompasses a pivotal role during the second and the third phase of a BTP and is defined as the process that runs in parallel with any other business process; assigning responsibility to people and technology on processing, capturing, maintaining and defining data accuracy based on a defined set of rules.

Multiple parameters relevant to MDM such as change management, no practical commitment from senior management, no compliance with any data governance policies, implementing new integrations or any pre-existing data quality challenges along with multiple others, can jeopardise the successful completion of a BTP. As MDM becomes significant in the second phase, the research focuses on how the invasive circumstances arising from such parameters during this BTP phase and beyond may be addressed by the BTP's programme directorate to enhance decision-making through the appropriate impact on MDM. The programme committee of a BTP would thus become aware of how to: a) manage master data, b) reinforce enterprise data quality and c) govern the overall BTP lifecycle by safeguarding data practices.

Alongside an extensive learned literature review and industry resources to establish the research aims from the outset, the research appropriated a deductive and interpretive research methodology to two Data audits as case studies plus a series of semi-structured interviews and subjected to a comprehensive qualitative analysis. Each BTP either faced challenges or was about to face challenges. The different roles of the participants and the different phases of each BTP in which the audits took place allowed the research to employ these multiple methods to reflect different aspects of the same issue.

Referring to the Data Audit Framework for added structure, the two data audits took place in two different companies. The first company was performing the audit after a failed BTP, and they had already an MDM function within the organisation. The audit focused on the performance of its already existing function. The second company had initiated a BTP and wanted to ensure that the required controls were in place for a successful delivery. These two audits provided valuable case study evidence for the evaluation of the decisions made during the BTP with regards to a) master data, b) what led the programme directorate to these decisions and c) how the decisions affected the outcome of the BTP as well as the organisation itself.

The interviews consisted of twenty-eight semi-structured questions and involved eighteen people with diverse backgrounds and from divergent functions of the business. All the interviewees were participating in a BTP with an underlying MDM process. The interviews provided evidence on a) how different roles within the programme reflect and react under specific circumstances and b) how each workstream prioritised data-related activities in conjunction with the overall programme.

From the case study audits and the interviews, the research identified an enhanced understanding of the reasons behind the decisions during a BTP concerning MDM, and how these decisions consequently affect the successful implementation of a BTP. From these findings, the research proposes a novel MDM-impacted BTP decision model that brings together its contributions to knowledge, and the basis for future work.

Preface

Being the Co-Founder of a Data-Driven Software company for over ten years now and acting as a Fractional Chief Data Officer or Data Architect on a consultancy basis in multiple Business and Digital Transformation Programmes since very early 2000s, I was very aware of how challenging was and is to divert and balance the emphasis of any programme's directorate to the data quality prior, during and after the implementations.

Ensuring Data Quality and Data Continuity is always a very important element of every Data Migration Strategy. Both data quality and data continuity highly depend on a clear understanding of the organisations static data or master data; that was mainly the challenging part of my job as a consultant.

The three components for every digital programme are people, process, and technology. Having to set up a strategy that highly depends on these three elements has been fascinating. The successful alignment of these three pieces is the key to successful delivery.

Every programme has high points and low points during its lifecycle; the reasons for my engaging as a consultant is because the lows overcome the highs, or because it is challenging to initiate the data element of the programme.

Even though this part was very challenging, at the same time it was very rewarding because it allowed me to understand different industries, different businesses, different processes, different technologies, different roles, different people and different behaviours on every project as a result to collect experimental data.

The research drew upon the opportunity to generate experimental data from multiple engagements during Business and Digital Transformation Programmes that was either happening at that time or had happened in the past. Or they were about to start and they were on the planning phase. These engagements allowed me to capture and understand the reasons that these programmes are extremely challenging to reach the successful completion of the initial inspirations; but also, to collect and analyse the reasons that these challenges appear and how they are resolved and resulted in a contribution to knowledge.

Acknowledgements

I cannot thank enough Dr Simon Polovina and Dr Sohrab Moshiri for their guidance and support. Thank you both for all these out of hours support and understanding!
I would also like to express my very great appreciation to Dr Keith Burley for his continuous support over the years. He was always there supporting me from my master's degree to this Research. Dr Burley encouraged me with his constructive and enthusiastic guidance. I would also like to thank Dr Paul Crowther for his support.
I would also like to offer my very special thanks to Mr Kevin Carrick CEO and Co-Founder of Data Clarity Ltd for his support and guidance over the years. Through his leadership, I have gained the experience that I have today.
Finally, I would like to thank my wife Anastasia Kyriakopoulou for her overall support and for the greatest gift that I could ever get, my daughter Kally-Nicole.

Dedication

To my daughter Kally-Nicole Lepenioti.

Contents

List of Figures

List of Tables

Abbreviations

Term	Definition
AX	Microsoft Dynamics AX
B2B	Business to Business
B2C	Business to Consumer
BI	Business Intelligence
BTP	Business Transformation Programme
CAPEX	Capital Expenditure
CFO	Chief Finance Officer
CIO	Chief Information Officer
CRM	Customer Relationship Management
CRUD	Create, Read, Update, Delete
DGO	Data Governance Officer
DTO	Data Take On
DW	Data Warehouse
EDS	Enterprise Data Strategy
EPOS	Electronic Point of Sale

ERP	Enterprise Resource Planning
ETL	Extract Transform and Load
MDM	Master Data Management
MIS	Management Information System
OPEX	Operational Expenditure
PIMS	Product Information Management System
POC	Proof of Concept
ROI	Return On Investment
SOA	Service-Oriented Architecture
UDM	Unified Data Model
WMS	Warehouse Management System

Glossary

Term	Definition
ACES	The name is given the SHU Faculty of Arts, Computing, Engineering and Sciences.
Bidirectional Integration	The process of integrating two or more business systems and to synchronise the data within these disparate systems to each other.
Big Bang Project	All the phases of the project in one release.
Big Data	Extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions.
Business Buy-In	The acceptance of and commitment to a specific concept or course of action by the senior management.
Business Intelligence	The term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance.
BizTalk	BizTalk is an industry initiative headed by Microsoft to promote Extensible Mark-up Language (XML) as the common data exchange.
Data Governance	DG is the overall management of the availability, usability, integrity and security of data used in an enterprise
Data Mining	A method of exploring data inductively.
Data Taxonomy	Taxonomy is the science of classification according to a pre-determined system with the resulting catalogue used to provide a conceptual framework for discussion, analysis, or information retrieval.

Data Warehouse	A large store of data accumulated from a wide range of sources within a company and used to guide management decisions.
False Decisions	Refers to all the decisions that the professionals have taken under pressure during the lifecycle of a project.
Metadata	A set of data that describes and gives information about other data.
Off the self Software	Off-the-shelf software is a mass-produced item that is adaptable to a wide variety of unique businesses.
Software as a Service	SaAS-IS a software distribution model in which a third-party provider hosts the application and makes them available to customers over the Internet.
Spreadmart	A set of Excel Spreadsheets maintained by a group of people to perform tasks.
SQL	Standard Queries Language for database interrogation
SQL-Server	Microsoft's Relational Database

Chapter 1: Introduction and Motivation for Research

1.1. Introduction

This research aims to identify the importance of Master Data Management (MDM) during a Business Transformation Programme (BTP) and how any decisions were taken during the lifecycle of the programme affect its successful completion and the organisation, overall.

Master Data Management as a process is extremely important and vital for any organisation. The understanding of the vital core information that exists within the organisation by the people who run these organisations is necessary to achieve operational excellence and accurate insights.

There are several definitions of what Master Data Management is. According to Berson and Dubov (2011), MDM is defined as: "*the framework of processes and technologies aimed at creating and maintaining an authoritative, reliable, sustainable, accurate and secure data environment that presents a single version of the truth for master data and its relationships, as well as an accepted benchmark used within an enterprise as well as across enterprises and spanning a diverse set of application systems, lines of business, channels, and user communities*" (Berson and Dubov 2011).

According to Gartner Group "*Master data is the official, Consistent set of identifiers, extend attributes and hierarchies of the enterprise*" and *"Master Data Management is the workflow process in which business and IT work together to ensure the uniformity, accuracy, stewardship and accountability of the enterprise's official, shared information assets*" (White and Radcliffe 2008).

This chapter introduces the motivation for the research and establishes the research aim. The research approach is outlined, using data audits and semi-structured interviews with professionals during their participation within a BTP.

1.2. Motivation

The motivation for this research came from being involved, reviewing, and evaluating the underlying data strategy in multiple BTPs. The involvement with multiple BTPs, especially during the planning and during the evaluation phase, resulted in common themes around the importance of data quality and data governance prior, during and after the programme's implementation. Businesses are now more advanced in their use of data, which drives more demands and more complicated ways of using this data. As stated by Al-Ruithe, Benkhelifa and Hamed (2017) there were many unsuccessful attempts in governing data as these attempts were mainly driven by IT. Data Governance and MDM remain informal with very ambiguous and generic regulations, isolated in siloes, lacking structure and wider support of the organisation (Al-Ruithe et al 2017).

In a different article Al-Ruithe and Benkhelifa (2017) suggested that the area of data governance, in general, is under-researched and not widely practised by organisations. Besides, Weber, Otto and Osterle (2009) tried to identify the strategic and operational challenges that the companies face with their demand to high-quality corporate data. In their article, they describe a data governance contingency model and they demonstrate the influence of performance strategy, diversification span, organisation structure, competitive strategy, degree of process harmonisation, degree of market regulation, and decision-making style on data governance. Based on their findings, companies can structure their specific data governance model (Weber et al 2009).

Data Governance is one of the key processes of MDM. Griffin (2006) stated that there is an incredible opportunity for organisations that by implementing an MDM solution to standardise their core corporate data across the enterprise; organisations will be in a position to put timely, consistent, accurate and actionable information in the hands of management and knowledge workers. However, one of the challenges that are stated in the article is the concern that the structure that is required to provide ongoing data governance, is too complicated or ineffective. It is a fact that data governance has its costs. However, ineffective governance is much more costly. Other challenges are the definition of the business value of the initiative as well as the reluctance to implement due to the bureaucracy that it creates

and the difficulties on prioritising the data standardisation process based on underlying business drivers. (Griffin 2006).

All these are part of a data strategy that should be a part of any BTP. Any Business or Digital Transformation Programme is divided into three different major phases. The first phase is the selection of the appropriate set of application systems as well as the introduction of new business processes across multiple lines of business and different channels. The second phase is the implementation of the new application systems and the data migration process. And the third and final phase is the transition from the legacy application systems and business processes to the newly defined framework of processes and technologies that ensure the business and data continuity. Master Data Management and it is Data Governance, encompasses a vital role during the second and the third phase of the programme and is defined as the process that runs in parallel with any other business process; assigning responsibility to people and technology on processing, capturing, maintaining and defining data accuracy based on a defined set of rules.

Based on all the above, there is a motivation to understand in more depth the reasons that these challenges exist. What are the actual drivers that professionals struggle to follow the best practices that are defined in the literature? And why is it so difficult to get the support and the general acceptance on any BTP, in which a robust data strategy is equally important with the overall programme? Moreover, successful implementation of the data strategy should be of similar significance as the overall programme's success and the programme directorate should consider them as of equal importance.". The motivation was even more intense in understanding at what point the programme overshadows the data strategy initiative and what are the reasons for this behaviour. Which roles are responsible for bypassing the framework that the MDM and data governance initiative has defined and set to follow, and when the support from the senior management is getting reduced and for what reasons?

With this research, there is a motivation to contribute to knowledge a small part on the reasons that there are data-related challenges during a BTP based on experimental data, and to enhance the literature as much as it is possible, so the practitioners and the people that are involved in BTPs to prepared on the challenges that they might have to overcome.

1.3. Research Aim

This research aims to establish the importance of the Master Data Management initiative during a BTP and to define the decisions and the circumstances were made during the programme and can affect its successful implementation. This is valuable to any future MDM initiative as well as any BTP.

1.3.1. Research Objectives:

The following objectives are explored within this research:

1. Provision of focus through the acquisition of knowledge of the reasons that cause data quality problems within an organisation.
2. To focus on the existing enterprise data quality and data governance methodology that is part of a BTP and the extent that is taken into consideration by the people that are involved in the programme.
3. Provision of focus through the knowledge representation of the MDM initiative and clarify the support to specific design areas during the planning of the MDM.
4. To enable knowledge sharing and reuse to practitioners to deepen the understanding of the challenges that may have to overcome during a BTP.
5. Provision of an understanding of the knowledge which is derived from the examination on how professionals in different roles respond to the importance of Master Data and Data Governance during the Business Transformation Programme and also, provision of an understanding on how BTP is affected by their perspective.
6. To provide an understanding of knowledge derived from how to approach people from different business functions during the BTP and how their commitment to the data governance framework can affect the BTP and the business, overall.

1.3.2. Research Question:

"*How may the invasive circumstances during a Business Transformation programme be addressed to enhance decision-making with its impact on Master Data Management"*?

1.3.3. Research Proposition Specification

Following the declaration of the Research Objectives and the Research question, this research focuses on the elements that can affect an MDM initiative during a BTP and what can be done to identify any misalignments at an early stage. That is why, following the Research Question,

it is set to understand what are the invasive circumstances that can affect any decisions made regarding Master Data and MDM, and how these circumstances impacting the BTP and the Organisation overall.

Since a BTP depends on Business Processes and Process Flows that generate data, the data assets that any organisation holds should be in a position to serve the needs of each business function. That is why ensuring Data Quality should be a high priority for each BTP. With the first Objective, the research tries to understand how Data Quality problems occur and what are the reasons for Data Quality problems. By investigating the reasons behind the Data Quality problems, this research will try to identify how these problems affect any decisions during a BTP and if these problems turn into invasive circumstances that can affect any Programme.

Following the first Objective, the second objective tries to investigate the ways that the Data Quality can be ensured by investigating how organisations apply an enterprise Data Quality methodology and how this methodology is aligned with the overall Data Governance. The reason for this objective is to understand how the people that are involved in a BTP follow any methodology. Based on the way that they apply any methodology, the objective will try to investigate if any circumstances arise from not complying with these methodologies and how can affect any decisions related to BTP.

The third Objective investigates the different implementation techniques for an MDM process. By investigating the different MDM implementation options, there was a better understanding of the expectations that people that participated in a BTP had. By realising the expectations that people have in comparison with what they have planned in their programme, can create multiple circumstances for decisions that can be characterised invasive.

Based on the first three objectives, the fourth objective tries to assist the practitioners in understanding when circumstances can be characterised invasive and in general can affect a BTP. The fourth Objective tries through the research to identify potential warning signs that may help the practitioners to overcome during a BTP.

The fifth and sixth objective tries to understand how people in specific roles and people in specific business functions understand the importance of MDM and Data Governance during a BTP. The research tries to understand the way that a business function or a specific role responds to any data-related tasks, and how other variables can affect their response to the needed attention and support that a data task requires. Based on these responses, these two objectives try to identify any circumstances that can be characterised invasive with regards to MDM and can affect the BTP and the organisation overall.

1.4. Research Overview

Alongside an extensive learned literature review and industry resources to establish the research aims from the outset, the research appropriated a deductive and interpretive research methodology to two Data audits as case studies plus a series of semi-structured interviews and subjected to a comprehensive qualitative analysis. Each BTP either faced challenges or was about to face challenges. The different roles of the participants and the different phases of each BTP in which the audits took place allowed the research to employ these multiple methods to reflect different aspects of the same issue.
A detailed presentation of the research methodology and the reasons why the specific research philosophy and approach have been selected can be found in Chapter 4 of this thesis.

1.5. Contributions to Knowledge Overview

The conclusion to this Research was the generation of three artefacts that can be characterised as invasive circumstances that affect MDM and consequently the BTP. These three artefacts are the following:

- This research contributed to the body of Knowledge by identifying some of the invasive circumstances that can affect the MDM and consequently the BTP and their potential impact in multiple stages.
- Also, following the theoretical best practices this research identified potential critical areas that may jeopardise the programme if these areas got overlooked.

- And finally, after merging the theoretical guideline and the practical implication, has been identified that even if the programme directorate follows the best practices, the timing and sequence of executing these tasks may affect the successful implementation and create more invasive circumstances.

All three artefacts consist of a decision topic. All these topics have been derived from the research objectives and each objective is aligned with more than one topic.

- Topic 1 has been formed based on the outcome of the research objectives 1,4 and 5.
- Topic 2 has been formed based on the outcome of the research objectives 2,4, and 5.
- Topic 3 has been formed based on the outcome of the research objectives 3 and 6.

All three topics together propose an MDM-impacted BTP decision model. This model offers an enhanced understanding of the reasons behind the decisions during a BTP concerning MDM, and how these decisions consequently affect the successful implementation of a BTP.

1.6. Overview of Thesis

This thesis, including this chapter, consists of six different chapters.

Chapter 2 consists of two sections. The first section starts by showcasing the MDM roadmap and framework that Infotech Research Group suggested (2014). Both audits and the Interviews had this framework as the baseline. The structure of this chapter explains each box of the roadmap separately. The second section provides the theoretical foundations of the MDM used in this research. The reasons behind the data quality problems that organisations face and a deep investigation into the methodologies of Enterprise Data Quality and Data Governance. It continues with an investigation of the importance of data quality. The chapter continues with the design areas when planning an MDM initiative and the overall MDM architecture and then dives into the worst practices and what needs to be avoided when designing the data governance process. The chapter finishes with an investigation on how data governance needs to be non-invasive.

Chapter 3 provides a detailed presentation of the Research Methodology and is consisted of four sections. The first section is the introduction which gives an overview of the chapter. The second section focuses on the research methodology. The research methodology consists of the research philosophy, the research approach and strategy, the primary and secondary research the feasibility and the ethics. The third section focuses on the research strategy and starts by presenting the research positioning. It continues with the design of the Audits that took place and generated experimental data and the design of the semi-structured interviews. The fourth section of this chapter is the concluding section and presents a critical reflection and research methodology and design.

Chapter 4 provides a presentation of the experimental data that collected by the audits and the interviews. The data audits that took place in two different organisations in the United Kingdom and the twenty-eight interview questions that have been performed with eighteen people with different managerial and technical roles as well as project management roles. All interviewees are heavily involved in BTP. The data quality and the MDM as a process was vital for the succession of the projects. On the first part, there is an explanation of the approach that was followed to perform the audits and the reason this approach had been selected back to the time that these audits took place. On the second part, the twenty-eight questions are analysed, explaining the reasons these questions were selected and how they reflect the original objectives and research question that has been defining on the introductory part of the Thesis.

Chapter 5 provides an evaluation and the presentation of the results along with the findings. There are six main different sections. The first section is the introduction that presents an overview of the chapter. The second section is a comparison between the two companies that the two audits took place. The third section consists of a critical reflection on the interviews. For each group of questions, there is an agreements and disagreements between the groups subsections as well as an impact on MDM and BTP subsection. The fourth section summarises audit reports and interviews highlights. The fifth section is the summary and cross-case evaluation. There is a combined list of highlights between the data audits and the interviews,

followed by a comparison between the data audits and the interviews with the state-of-the-Art MDM Framework and Roadmap. The section concludes with an overall comparison of the state-of-the-Art MDM Framework and the theoretical foundations of MDM. The chapter finishes with the sixth and last section which is the contributions to knowledge based on the findings of this research.

Chapter 6 is the conclusions and further work. The chapter starts with the first section which is the contribution to knowledge. The second section is the thesis statement followed by the conclusion of this thesis. The Chapter finishes with an extended section focusing on the limitations of this research and any further work that can be done.

Chapter 2: Theoretical Foundations and State-of-the-Art MDM Framework and Roadmap

2.1. Introduction

This chapter is focused on understanding the best practices that are used when an MDM solution is implemented and evaluates the existing literature on understanding the foundations of MDM and what are the route causes that invasive circumstances can arise. The chapter consists of two major sections. The first section starts by showcasing the MDM roadmap and framework that Infotech Research Group suggested (2014). Both audits and the Interviews had this framework as the baseline. The structure of this chapter explains each box of the roadmap separately. The second section provides the theoretical foundations of the MDM used in this research. The reasons behind the data quality problems that organisations face and a deep investigation into the methodologies of Enterprise Data Quality and Data Governance. It continues with an investigation of the importance of data quality. The chapter continues with the design areas when planning an MDM initiative and the overall MDM architecture and then dives into the worst practices and what needs to be avoided when designing the data governance process. The section finishes with an investigation on how data governance needs to be non-invasive. The chapter finishes with an evaluation between the framework and the theoretical foundations and a table that captures key findings from both framework and literature. Each subsection finishes with a critical evaluation.

2.2. The Framework

During this research, there was a large involvement in implementing data strategy for new ERP implementations as well as applying MDM and data governance processes. (*This involvement includes thirteen BTPs advising on the data strategy and setting up data governance. Fourteen data related audits were also performed focusing on the way that data integrates between various systems and auditing the selected process that had been applied in each case. There were also three failed BTPs.).*

Based on this experience, an approach to how an MDM process can be implemented had been followed. By investigating multiple options on what is the most appropriate way to initiate an MDM implementation, a graphical representation of an MDM roadmap that includes all the steps that need to be taken into consideration as the first step. This graphical representation was initially published by Infotech Research Group in March 2014. This graphical representation is presented in figure 2, as the foundation for any MDM implementation.

This Chapter analyses and investigates the MDM roadmap and framework that Infotech Research Group suggested. Most of the involvement that is mentioned in the previous paragraph followed this framework to achieve successful results. The structure of this chapter explains each box of figure 2 separately.

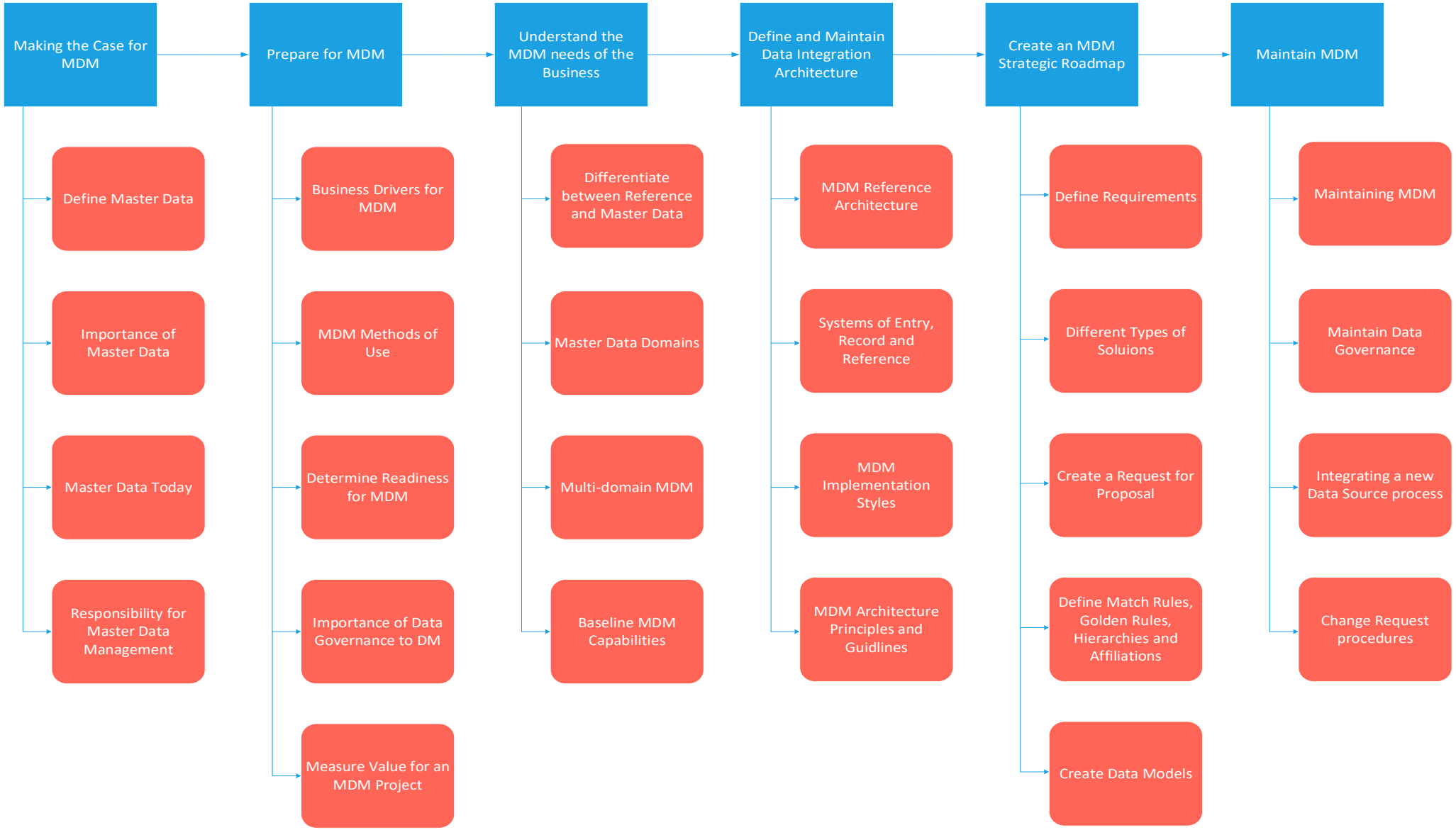


Figure 1 MDM Implementation Roadmap: Infotech Research Group 2014

2.2.1. The case for Master Data Management

The purpose of this section is to make sure that the team who is setting up the process of MDM and will be responsible for the implementation from a technical perspective, can clearly explain the importance of the process.

2.2.1.1. Why Master Data Management is essential

Any organisation in which the data is shared across multiple systems has the responsibility and the necessity to have MDM in place. An MDM solution will provide the organisation with a single accurate set of data populated across multiple systems.

Usually, the people that are involved in a project like this are all part of IT, however, MDM is a process that is depending and affecting many business functions. That is why, what is master data and why master data is required must be defined. As part of this definition, it is important to have a clear definition of the goals of a programme like MDM.

a) What is Master Data?

Master data addresses critical business entities that fall into four broad groupings. These four groups are people (customers, suppliers, employees etc.), places (physical spaces and segmentations) things (products, services) and abstract concepts (hierarchies used for reporting, accounting etc.). This data is typically critical to the organisation, less volatile, more complex, contains many data elements (attributes) and is used across multiple systems (InfoTech Research Group 2014).

b) Why do you need Master Data?

While datasets are often used for different purposes, the same types of data are often shared across the organisation. Master data is the accurate set of data that is used across the organisation and populated in multiple systems. That is why this consistency within the systems

is required so, everybody in every business function can understand their role that they have to work within any case (InfoTech Research Group 2014).

c) What is Master Data Management?

MDM will detect and declare relationships between data, resolve duplicate records and make data available to the people processes and applications that need it (InfoTech Research Group 2014).

2.2.1.2. Benefits of MDM

The purpose of MDM is to ensure that everyone in the organisation needs to realise the benefits that they will get from their point of view. That is why there is a requirement to establish a set of benefits that will be understandable from IT users as well as business users. Both user types want to use the same reliable dataset, but the benefits that each one of them will get is different based on each user's perspective (InfoTech Research Group 2014).

a) Business Benefits.

The importance of having an MDM solution in place and as well as the definitions of both master data and MDM has already been given. However, this process applies to business functions as well as IT functions. For that reason, a presentation of the business benefits is as follows:

- Increasing operational efficiency by eliminating the need to reconcile the same data from various sources (InfoTech Research Group 2014).
- Make better decisions and gain deeper insight (more consistent reporting and better visibility into operations that helps businesses to make decisions). Some people might mention that the insights and the decisions that the business will take can come from the business intelligence platform. However, people do not seem to realise that all the insights come from the data that have been collected from multiple sources. If there is no consistency on all the multiple sources, then there will be no consistent results to make decisions on (InfoTech Research Group 2014).

- Comply with Complex regulations (InfoTech Research Group 2014).
- Increase customer satisfaction by enabling the organisation to access a 360-degree view of the customer data (InfoTech Research Group 2014).
- Easily make and manage changes to master data entities as business needs change (InfoTech Research Group 2014).

b) IT Benefits

MDM cannot help IT only by ensuring that the data is accurate and consistent across the systems. Ensuring data integrity within the data warehouse which is the repository for any business intelligence tool by maintaining a dimension management process, is not the only benefit that MDM will offer. The issues that an MDM solution will benefit IT are as follows:

- Drastically reduce the number of issues and incidents reported by the business regarding untrustworthy data (InfoTech Research Group 2014).
- Improve data quality across the key systems. This data quality improvement will be expanded on other systems as well since the integration between primary and secondary systems will be based on a qualitative basis (InfoTech Research Group 2014).
- Reduce the blame IT receives for poor data quality (InfoTech Research Group 2014).
- Prepare for a data audit and ensure compliance with complex regulations (InfoTech Research Group 2014).
- Gain insight into where master data sits within the organisation's system environment (InfoTech Research Group 2014).
- Understand the business and address business needs (InfoTech Research Group 2014).

d) What is the goal of Master Data Management?

The end goal of an MDM implementation is to make sure that the organisation's investment in this technology will deliver the promised business results by supplementing the technology with rules, guidelines, and standards around the enterprise data. This will ensure that data will

continue to be synchronised across data sources on an ongoing basis (InfoTech Research Group 2014).

Not all data that falls within the groups of people, things and abstract concepts need to be managed as master data. MDM is not “one size fits all” (Weber et al 2009) solution. During the MDM strategy development, the organisation must identify what data is critical and should be managed as master data (InfoTech Research Group 2014).

2.2.1.3. Importance of Master Data Management

MDM becomes increasingly important with the rise of cloud computing, business intelligence and big data. Based on that, MDM will not only help manage and make sense of the large amounts of data being produced and consumed by the business but will also help in the analysis of this data. These are the main reasons why MDM is important:

- The increasing volume of data and the use of analytics can produce more actionable and useful information. However, many of these data sources are fragmented and inconsistent. MDM can help remedy this by pulling information together and fixing quality issues (InfoTech Research Group 2014).
- It can also help the business work around big data by helping manage and make sense of the large amounts of data being generated. This can happen with the appropriate classification and taxonomy of a large amount of incoming unstructured data (InfoTech Research Group 2014).
- Following the above, MDM can create a meaningful way of indexing big data. If the business is aware of what master data exists within it, then it will be aware of which data is most critical. This means that when the business begins looking at big data, then it can quickly target only the data that is critical and is required (InfoTech Research Group 2014).
- Also, with the social media revolution, the business could gain a full view of its customers. With proper integrations between customers and their social media, the business can get

a full 360-degree view of attributes that the customer can be associated with a segment (InfoTech Research Group 2014).

2.2.1.4. Who owns the responsibility?

As previously described, MDMis a process that involves both IT and Business. The fact that it is a process, the implementation requires input from both parties. IT should be responsible to select the tool and to implement the solution with all the required integrations. And the business should define the process that will be affected and how this will improve the way that everybody works in the medium and long term. Therefore, when master data is created, the business needs to make sure that their input and dedication is confirmed to ensure that IT can make decisions that fit with the corporate goals and objectives. Thus, the following three points ensure balanced cooperation between IT and Business.

- First, the data belongs to the entire organisation without separation between IT and Business. Therefore, master data is a responsibility for everybody to ensure that the creation of master data is happening based on the needs of the entire organisation.
- The ownership of the master data is the responsibility of the business. In this case, IT is responsible for the project's technology, support, platforms, and infrastructure. However, the ownership of the business rules and standards lies with the business. IT cannot take the responsibility of defining the business rules since there is a danger that the logic, focus, and state of mind between the two sectors are completely dedicated to several aspects.
- Therefore, IT and Business need to form a partnership to implement MDM as a process and as a solution. As a result, IT is responsible for the technical component while the business will be the key in identifying the master data.

Subsequently, a representation of the above into a responsibility matrix would look like the table below:

Responsibilities	**IT**	**Business**
MDM Requirements	√	√

MDM Business Rules		√
MDM Technology	√	
MDM Support	√	
MDM Infrastructure	√	
MDM Standards		√

Table 1 MDM responsibility matrix(InfoTech Research Group 2014).

With this level of cooperation, there is always a challenge that the MDM project will probably have to face. This challenge is the alignment of the multiple departments to work together to achieve the same long-term vision. Before a data repository build and design, there must be a confirmation that the same level of understanding between IT and business, as well as the expectations and the requirements, are aligned. There must be a level of engagement between the IT and the business users, so their specific current and future needs, requirements and expectations are met and can be solicited and incorporated within the MDM plan.

Critical Evaluation

The MDM initiative is a process that affects the entire organisation. As per the above table, there are specific elements that can only be done by the IT function of the organisation, and some that can only be done by the Business function of the organisation. However, MDM is a combination of Business, Process and Technology and all these three elements need to be fully aligned with a very clear and understandable action plan. There is an inter-dependency between the functions and each one of them should support and supplement each other. It is of high importance that people in these different functions do not just assume that this task is not my functions responsibility and vice versa. Every part should be sharing the same vision.

2.2.1.5. Implementation Approach

After the definition of master data and MDM and having defined the benefits of the process and declared responsibility for both IT and business. The next step is the definition of the implementation approach for the MDM.

Based on the nature of the organisation and the level of complexity within its IT systems, the data domains that an MDM implementation should focus on must be declared.

According to the above statement, considering a complicated technical eco-system or not, the right approach would be to define all the architectural decisions and tool choices. These decisions should be made with all possible master data entities in mind and to implement the process through a series of on-going, smaller projects rather than a onetime “Big Bang” project. Therefore, the project:

- Should be guided by the organisation’s architecture, business priorities and roadmap.
- Should focus on one data domain area at a time. This would make it easier to demonstrate the value of MDM quickly. A data domain in most of the cases may consist of customers, vendors, products, and locations.
- Should drive the organisation to consider all data it may want to be mastered while the project should focus on one mastered data domain. Architectural decisions and tool choices should be made with this holistic picture in mind. This would prevent choosing a solution that is too specific which would result in the need to purchase additional solutions in the future to support other data domains.

2.2.2. Getting Prepared for Master Data Management

Before the next step which is the establishing of the business drivers for a project like MDM, it is essential to ensure the buy-in from the business. This action will ensure that the MDM vision is aligned with the business drivers.

For that reason, it is essential to make clear, how the IT part of MDM benefits can help the business on understanding the success of an implementation.

In most cases, master data can be perceived as an IT-related issue. However, MDM exists to enable the success of an organisation on its entirety. A successful implementation will become the lifeblood of the organisation and will help all facets of the business.

2.2.2.1. Business Drivers for MDM

The business drivers for MDM can include:

- Increased revenue: Depending on the method of use, a successful implementation can provide the business with the ability to cross-sell/up-sell, customise offerings, improve customer retention and introduce new products.
- Agility: The organisation can consolidate data from silos and integrate it into new systems. This allows the business to identify relationships and hierarchies and adapt data to the changing needs of the business.
- Cost: Manual business process can also be automated. Efficient and accurate invoicing and a consolidated ERP and CRM system.
- Compliance: An MDM system will reduce the risk, ensure data adheres regulations, and ensure customer privacy preference

2.2.2.2. Master Data Management Methods of use.

As part of any implementation, a declaration of how the MDM is going to be used is necessary. Most importantly, each organisation will have to understand what the expectations will be once the process is in place. For that reason, important questions that answer the where, the how and the who need to be clear from the beginning.

- **Where:** The where method of use defines the systems that the MDM will be applied. What type of systems is currently being involved in the collection of master data? Do these systems process master data in real-time? What is the importance of capturing the master data on the creation? To answer all these questions, the definition of the operational level of the system needs to be defined and based on the outcome, and the strategy will have the foundations defined and ready to move to the next step. This method is the **collaborative method of use** (Obenhofer and Dreibelbis 2008) (Chapter 3.4).

- **How:** The how method of use defines the type of data that are part of the MDM Process and how this data is going to be used. For example, does the organisation needs master data to be a part of business intelligence so the insights will drive the decision-makers to more accurate decisions? Does the organisation want to perform any identity analysis to identify any unclear related transactions between its operational transactions? Or the organisation just requires a clear view of how many customers got introduced to the business over a period? This method is the **analytical method of use** (Obenhofer and Dreibelbis 2008) (Chapter 3.4).
- **Who:** The who method of use defines the people that are going to be involved in the MDM and how these people are going to work together. These people can be in different roles and different departments, but they all get involved in the same business process. For example, when a new product is created, many teams are part of this process. These teams need to work together and to be aligned to complete this process without overlapping or overruling each other. This method is the **collaborative method of use.** (Obenhofer and Dreibelbis 2008) (Chapter 3.4).

2.2.2.3. Determine the Readiness for MDM.

A maturity assessment is required for determining the readiness for MDM. This maturity assessment is used to identify potential areas within the organisation that need development. During the maturity assessment, the focus is on five different elements that are required for the MDM process. These are the current architecture, current data, current processes, current management/administration, and the current governance. These five different elements can have up to five different stages of readiness. These stages are the initial, the fragmented, the standardised, the integrated, and the optimised.

	Initial	Fragmented	Standardised	Integrated	Optimised

Governance	Responsibilities not assigned.	Driven by IT. No business Implication.	Governance in each domain	Cross-domain governance	Global governance for all the data of the company
Management/ Administration	No data administrators.	A data management culture is being spread within the organisation	Teams are in place but on a limited scope	Centralisation of the federation of data admins	The management of the static data is part of the organisation's operational processes
Processes	No interfaces between applications or functions.	Initiatives begin to consider data lifecycle	Focus on managing data is limited to one domain	Best practices are shared across the organisation	Continuous improvement of the data lifecycle
Data	No data quality metrics.	Few data quality metrics	Metrics are used at the domain level	Performance is monitored with cross-domain metrics	Performance is monitored with cross-domain metrics and this monitor is part of the organisation's operations
Architecture	No tools for managing data quality.	Existing tools for managing Data Quality but no MDM system	1 or multiple MDM Systems	MDM systems are consisted and integrated	MDM systems are consisted and integrated into a synchronous mode.

Table 2. Maturity assessment for MDM (InfoTech Research Group 2014).

Critical Evaluation

The Data Maturity assessment can be characterised as one of the most important tasks of the MDM framework. The entire programme should depend on an understanding of the organisation's data. If there is not a clear understanding on where the data is, what the data means, how the data is used and what are the limitations of accessing this data and how this data can be integrated with the new system requirements, then the entire programme can be built its planning and architectural foundations in inaccurate assumptions. This inaccuracy can lead to increased costs on resources and services and important delays with third party suppliers.

2.2.2.4. Importance of Data Governance to Master Data Management

In every organisation, that there is a transformation process, the people that are involved need to follow a specific set of rules, so everybody can understand what liberties they have and what

liberties they do not have. That is why the following table shows how the governance rules apply to each function, what elements of the functions need to be under the governance, and why.

Governance for organisation culture	
What	**Why**
Sponsoring	Senior sponsorship of MDM initiatives to be built into processes, procedures, and organisation structure to ensure MDM gets the necessary attention across the organisation
Process and procedure framework	Data standards, data definitions and process escalation to make design decisions and ensure control of the processes in BAU (show how all the processes joined together)
Change management	To provide readiness for tool and organisation changes
Roles and responsibility matrix	To map tasks to the employee job role and ensure accountability
Training and Communication Plan	To provide readiness for training and raise awareness
Robust Accessibility Controls	Provides the systematic controls to prevent unauthorised changes
Governance for data and technology	
What	**Why**
Data quality framework and definitions	The rules and mechanisms to maintain data quality (availability, accuracy, consistency, reusability, redundancy)
Tools to manage the data quality	Specific tools such as data profiling help report on and track data issues against established business rules and standard checks
Tools for MDM	Required to enable and support the process of Create/Read/Update and Delete with the necessary controls and procedures to resolve bottlenecks and aid accuracy of data input
Governance for process	
What	**Why**
1. Process design for data creation and maintenance **2. Metadata and management of business rules** **3. Quality of solutions provided** **4. Audit and compliance** **5. Security and control**	For every field being maintained by the BAU MDM team, there needs to be a process defined with supporting business rules and process impacts of the data. This has an impact on the tool functional requirements, organisation design, data quality, training requirements – and is a foundational aspect of all MDM design and activities

Governance for Architecture and Standards	
What	**Why**
1. Data architecture per subject area **2. Data Model** **3. Data flow diagrams** **4. Strategy and integration of systems** **5. IT strategy and integration**	Data dependencies from a process or systems integration point of view have to be well understood to inform decisions around data loading, data design changes, data maintenance process changes and IT system strategy. If a system is due to be replaced or upgraded the cross-data impact needs to be defined and understood

Table 3. Importance of Data Governance (InfoTech Research Group 2014).

2.2.2.5. Measured Value for a Master Data Management Project

Before moving to the third stage, and after all the previous steps have been defined, it is important to assess the value that the project will bring to the organisation.

First, the organisation needs to evaluate again the reasons that they started such a project. Then to review the maturity assessment and where the problem lays with regards to governance. Once they have a comprehensive understanding of the above, the organisation will be looking for:

- Trust in master data. This trust will save hours of searching in multiple different systems trying to reconcile the data.
- Governance on data entry which raises more awareness to the users and changes the organisation's culture by establishing and following standards that ensures data integrity.
- With regards to ROI, it needs to be clear that at that stage, the organisation can measure the current processes that are in place. How much these processes cost to the business before the MDM implementation. Once the implementation been completed, the same metrics (usually these metrics can be taken during the maturity assessment if the organisation has not already measured it) can be applied and compared every annual quarter after the implementation.

2.2.3. Understand the MDM needs of the business

Each organisation has different needs, different processes, and different data. At this stage, the activity that takes place is the establishing of the standard actions which will set the high-level tasks for the MDM implementation.

2.2.3.1. Differentiation between Master and Reference Data.

There is usually a misunderstanding between master data and reference data. When an organisation decides to start an MDM implementation, they are struggling to understand which the actual master data is. This difficulty usually leads to occasions that rules are being created for every single attribute that it is used within the data domain. As a result, this creates fragmentations and challenges from the first step.

The framework disagrees with organisations and sets as starting point small and achievable objectives/ milestones. The main objective is to make the business understand that master data is different from transactional data that affects a business process or explains the clarification of a data record. The first type is master data and the second type is reference data.

2.2.3.3. Multi-Domain Master Data Management

At this stage, the business needs to define all the domains that need to be a part of the MDM process. The proposed framework suggests one domain at a time. Usually, there is a lot of enthusiasm at the beginning of a project and the businesses want to do everything at once by assigning multiple teams working in parallel.

However, this does not always work. It is suggested that due to the new rules that the data governance will force, the users will have to adapt to the new way of doing their jobs in a controlled way. The changes in the business processes performed by the users will have to be

applied gradually. If suddenly there is a massive change on the way that the users operate a specific process, there will be an equivalent risk of noncompliance by the user.

It is always more successful to apply the new policies in a controlled way that can be absorbed by the users and then once the users are comfortable with the new way of things, to move to the next one.

2.2.3.4. Baseline of Master Data Management

The baseline requires the established management buy-in as well as the understanding of the current MDM environment. At this stage, the implementation has a clear understanding of the current environment and a vision for the next steps. They also have a satisfying vision of how the project will proceed during the next steps which involve more technical details.

2.2.3.2. Master Data Domains

At that stage, the process becomes more technical. Having a data domain defined, it means that the process starts having a different level of objectives. As has been mentioned, each organisation has different requirements and different types of data. That is why the data domains that each implementation focuses can be varied from business to business. The data domain is the collection of all the attributes and subsets of data that defines a set of data.

The most common domains for each organisation are the set of groups that define a person (physical or legal) like customers, suppliers and employees, and the set of groups that define things like products, components, items, vehicles etc.

2.2.4. Define and Maintain Data Integration Architecture

Before the actual development starts, some technical requirements need to be clarified. These requirements are the following:

2.2.4.1. Master Data Management Implementation styles.

Most organisations manage their master data within multiple different systems within the enterprise. This reduces the data quality as well as the governance over the action of what is created, when, why but most importantly how this data is distributed across the multiple systems that the enterprises run their processes. The MDM system is going to be in place to control this distribution of master data within the organisation. That is why there are three different styles to define these integrations of data as well as to label the implementation.

- The first implementation style is the **Registry Style**. In which the MDM solution works at the back end collecting all the master data that is created, edited, or deleted on the operational systems. Then the users can go to the solution and review what has been modified or what duplicate records have been created. This is usually the first and most cost-effective style (Obenhofer and Dreibelbis 2008).
    - This style is usually recommended as phase one of the delivery of the MDM implementation project instead of going on a “Big Bang” approach. The suggestion is based on the level of data maturity that the organisation has been classified during the maturity assessment.
    - It is believed that a successful project, is a project that delivers visible results fast, by reaching quick and successful milestones.
    - MDM is a very difficult project that involves the biggest part of the organisation to be involved. The lack of quick visible results always causes disruptions on the project. That is why, setting expectations in multiple levels, turns to be more beneficial for the final goal.
    - One of the parameters that need to be taken into consideration is the budgets that are dedicated to the MDM solution. That means that usually there is not a

budget for applying changes to the existing systems (ERP, WMSs, etc) that have been provided by a third vendor. As a result, expectations should be clearly defined and documented. If the expectations are not correctly set, there is a risk to face budget challenges and constraints based on the first presentation of the outcome.

- The second implementation style is the **co-existence style**. This style is based on the fact that both operational systems and the MDM system are performing MDM processes. However, the golden record management is not happening in real-time but in an asynchronous mode. That means that the system captures the data, distributes the data to the MDM system, the MDM system performs the data quality and data completeness validations and then sends the data back to the application which then processes them (Obenhofer and Dreibelbis 2008).
- The third implementation style is the **transactional style**. This style is based on the fact that both operational systems and the MDM solution are performing MDM processes with the difference that the integration and the distribution are happening in real-time. What happens is that the operational system takes the data from the input, process them as normal and then the services from the MDM system validate the modified information and if there is a need for overwriting, the solutions act autonomously (Obenhofer and Dreibelbis 2008) (Chapter 3.4).
    - This is always the desired level of each implementation, but this is also the most expensive. It is always better to set more feasible and cost-effective targets for the deliverables at the beginning and then based on the results to set more complicated and expensive targets. Once the first benefits are visible, the business will understand the value of the deliverables and the benefits. Then the budget allocation is less challenging allowing the project the time that is required since there is already a solution there that shows improvement from the previous state.

2.2.4.2. Master Data Management Reference Architecture.

The MDM reference architecture supports the implementation of any of the MDM methods of use with any of the implementation styles that described previously delivering business solutions that supports MDM capabilities. This reference architecture is the best in class solutions that have been implemented as will be mentioned in the interview questions. The reference architecture is the sharing of knowledge, patterns and best practices from organisations that have already implemented an MDM solution and faced problems during the implementation. The reference architecture is the best practices of the MDM implementation. That is why no matter what organisation is going to implement a solution, they should always use the best practices and applies these practices to the needs of the organisation. The benefits of following what others have done successfully in the past help the project in the following aspects.

- Following an already successful architecture reduces the risk of the current implementation.
- The solution design does not need to start from the beginning as the main part of the designs has a reference back to the best practices that have been defined.
- The business can make more comfortably any decisions when there is a reference to back the proposal.
- The deployment requires less time as long as the project has a baseline that has been proven successful.

Critical Evaluation

It is very important to select the appropriate implementation style that can be supported by the organisation. As it will be investigated in more detail later in this chapter, one of the reasons that an MDM initiative could fail is due to the high expectations and confidence of what can be achieved by the programme without taking into consideration parameters such as, data maturity assessment, technology and system capabilities, and Data Governance compliance. Senior management has an idea of the technicalities that form MDM but they are not and they should

not be fully aware of the details. However, the expectations on how the MDM process works should be very clear to all levels. There are multiple instances that senior management and the programme directorate believes that at the end of the programme, the organisation will have a transactional style MDM while in reality, the system and the overall foundations of the organisations can support only the registry style implementation. That is not clear to the senior management as a result when they realise that the outcome does not match the expectations, the business needs to adjust to the new reality.

2.2.5. Create a Master Data Management Strategic Roadmap.

2.2.5.1. Define Technical Requirements.

At this stage, all the technical requirements have been documented. All the systems that are involved with all the technical details and the integration details are documented. All the data dictionaries, the data definitions, and the data normalisation between the systems with regards to master data have been produced. Based on the analysis and the documentation, the configuration workflow documents are produced and based on that, a plan for the development is ready to be initiated.

2.2.5.2. Different Types of Solutions

Based on the requirements of the organisation as well as the type of the implementation the organisation has already done the due diligence on the external vendors or an in house developed master data solution. As it is stated in the interview questions, most of the organisations select the option of an off-the-shelf solution.

- If the solution is an in-house developed solution, then all the resources that will be involved in the project will need to be allocated to tasks based on the internal project plan.
- If the solution is an off-the-shelf solution, all the agreements with the vendor as well as the internal resources that will be absorbed by the project need to be allocated on the project plan.

2.2.5.3. Define Match Rules, Golden Rules, Hierarchies, and affiliations

At this stage, the business has completed all the matching rules for the definition of the golden record as well as the required attributes that define a master data record. Hierarchies and affiliations for the master data have been defined and have been documented. These rules will be used on the data governance policy as well.

2.2.5.4. Create a Request for Proposal

At this stage, everything is ready and has been agreed. All the required documentation is there and a request for proposal is ready to be raised as a confirmation and a binding that these are the rules that define the MDM and the data governance policy. Based on this agreement, the actual development, if the approach is in house, starts or the configuration and the development of integration start if the approach is off the shelf solution.

2.2.5.5. Create Data Models

Once the request for proposal has been agreed and signed, the creation of the data models and the integrations starts. The development tasks will follow the agreed solution that is described on the request of the proposal.

2.2.5.6. Produce Strategic Roadmap

Once the implementation has been completed from a development point of view, strategic roadmap documentation needs to be created based on the solution that has been delivered. This roadmap will define the use of the system, maintenance of the system and general management that will be defined on the next and last stage of the MDM implementation lifecycle.

2.2.4.3. Master Data Management Architectural Principles.

Each architecture needs to follow specific rules to be successful. Once these rules are defined, the project team knows what they can and what they cannot do. These principles can be categorised to the following:

- The MDM solution should be able to get the data from the operational systems. Process this data and transform this data into information and asset of the organisation.
- The MDM solution should ensure data integrity and apply controls set by the organisation with regards to the integration of master data across the organisation in a consistent way. It should reduce the time and the cost of the data management and improve the quality and the comprehensiveness of the data.
- The MDM solution should be flexible on changes that can happen on master data if the business decides to change the process.
- The MDM solution should maintain the ownership of the data from its creation to decommissioning.
- The MDM solution should be flexible to connect with all the systems that the organisation needs to include in the MDM process.

2.2.4.4. Systems of Entry, Record and Reference.

At this point, the implementation has already defined the case for MDM, the organisation is ready, and they understand the needs. With the definition of the integration architecture, the implementation is ready to go into the analysis of the systems that need to be involved within the MDM environment. At this stage, all the systems that are going to be involved in the integration are documented and the data normalisation, as well as the data definitions and dictionaries of the attributes that exist on each system, are defined. The implementation is ready to move to the next step.

2.2.6. Maintain Master Data Management

2.2.6.1. Master Data Management Maintenance.

In this section, the team who delivered the solution needs to provide documentation with regards to technical maintenance. This documentation should include

- Infrastructure details
- Database administration details
- Integration details
- System user guides
- Training procedures for newcomers
- System troubleshooting and SLAs

2.2.6.2. Maintain Data Governance

The next documentation should focus on data governance. How new rules will be applied, how the users will be trained on the policy and details on how they will use the integrated systems with regards to data management.

2.2.6.3. Integrating a new Data Source

The next documentation should focus on the rules that already exist with regards to integrations with other data sources. The new data source should follow the rules as long as there is a requirement to be included on the MDM ecosystem. All the relevant documentation should get an updated version and all the data dictionaries, data definitions and data normalisation definitions should include this new source. The data governance policies should be updated as well to include the new source.

2.2.6.4. Change Request procedures

The last part of the MDM maintenance should be focused on the change request procedures. Any change that will happen to MDM should have a reason and this change should be documented. Any updates on the system or any updates on the integrated systems should be approved by all the involved parties, for example, if there is a change on an ERP system, the MDM team should review this change request and vice versa.

Critical Evaluation

Successful delivery of an MDM project is not only measured on the fact that the project a. meets the expectations and the objectives, b. finished on time and budget, c. there is a clear ROI, but it is also measured on how the process performs after the project lifecycle. MDM and Data Governance is an ongoing process with policies and rules that should be followed at all times. It should also be flexible and adaptable to new business process changes without the need to require a new project every time that a change needs to take place. A clear roadmap for the data lifecycle should be in place and should be easily used at the creation of new data, management of existing data and decommissioning data at its end of life (ie. GDPR etc). A successful implementation can be characterised by the longevity of its process.

2.3. Theoretical Foundations for MDM

This chapter focuses on the reasons behind the data quality problems that organisations face and a deep investigation into the methodologies of Enterprise Data Quality and Data Governance. It continues with an investigation of the importance of data quality. The chapter continues with the design areas when planning an MDM initiative and the overall MDM architecture and then dives into the worst practices and what needs to be avoided when designing the data governance process. The chapter finishes with an investigation on how data governance needs to be non-invasive.

2.3.1. Reasons for Data Quality Problems

As Maydanchick (2007) states one of the major problems that companies face is disrupted data quality which mainly results from necessary changes in the database, the business rules, and the user interface. Poor data quality is responsible for operational inefficiency, incorrect conclusions that lead to constrained decision making which finally leads to customer attrition. Knowledge management is used to overcome and to minimise the poor-quality data as it is the strategic process of integrating the new information that a company uses with the information that is stored in the system and to use them effectively. However, the specific task is not as easy as many vital company procedures are responsible for affecting the reliability of the information.

To shed light on the causes of data quality problems, the processes that are mainly responsible for disrupting the information are divided into three different groups: the processes that bring data from outside the database, the processes that use data that are handled inside the database, and the data that already exists in the database but becomes obsolete mainly due to time, and not due to changes in the database. Concerning the first group, it is claimed (Maydanchick 2007) that data conversion is the major source of problems in the database. The

data quality problem arises because in the data conversion period there is a lack of correct time allocation to three distinct layers of every system: database, business rules and user interface. Also, data consolidation is a major source of data quality problems and must be treated with attention as there are numerous data conflicts in the different database that are merged. Besides, it should be stressed that the common problems that arise from the manual data entries, the batch feeds and the real-time interfaces. Regarding the processes that change the data within the already existed database, Maydanchick (2007) claimed that data processing, data cleansing and data purging are responsible for distorting the quality of the data. While data processing has always been a source for potential problems in the database, recently data cleansing has gained increasing attention as a very careful methodology should be applied to avoid problems in the database. Finally, data that are already in the database may create problems (the so-called data decay) due to changes that are not correctly captured, there are upgrades in the database system or process automation that lacks validation. Hence it is suggested by Maydanchick (2007), that data quality management should be in use to minimise the consequences of distorted data.

Following the above, Maydanchick (2007) has identified thirteen different causes for data quality problems associated with the three different groups as follows:

- Processes Bringing Data from Outside the database
    - Initial Data Conversion
- System Consolidations
    - Manual Data Entry
    - Batch Feeds
    - Real-time Interfaces
- Processes Changing Data from within
    - Data Processing
    - Data Cleansing
    - Data Purging
- Processes Causing Data Decay

- Changes not captured
- System upgrades
- New Data Uses
- Loss of expertise
- Process Automation

Following Maydanchick's (2007) approach, Sarsfield (2011) identifies multiple reasons for data quality problems. Nearly all of them align with Maydanchick (2007). However, in addition to the identification of the problems, Sarsfield (2011) comes with a plan for resolving these causes that are numerically listed below. For each cause number, the letter (**a**) describes the cause and the letter (**b**) describes the solution.

The initial and most common cause of data quality problems are:

1. Typographical Errors and Non-Conforming Data
    a. Users are always vulnerable to mistyping information on an entry form when there is not the required validation or when the entry field **does not clearly state** what information is required.
    b. The solution to this problem according to Sarsfield (2011) are four options. The three out of the four options match exactly with Guess (2011):
        i. **Training**. Sarsfield (2011) and Guess (2011) suggest that users who enter the incorrect information in these fields can understand the impact that they have on the application and the business processes.
        ii. **Metadata Definitions**. Both Sarsfield (2011) and Guess (2011) suggest that by forcing validations on the fields, the users know exactly what they have to enter following directions from the metadata that each validation shows to the users.
        iii. **Monitoring**. Both Sarsfield (2011) and Guess (2011) suggest that by publishing the results of inaccurate data and by rewarding those who entered the data correctly, an organisation can keep track of what and

who entered wrong information and who continues to enter wrong information.

iv. **Real-time validation**. Sarsfield (2011) suggests a solution that "provides the ability to deploy data quality in application server environments, in the cloud or an enterprise service bus. As described in chapter 2 under 2.1.4, Obenhofer (2008) and Dreibelbis (2008) describe this approach as Transactional Style MDM Implementation, in which the MDM solution is integrating with the applications in real-time and can apply the MDM policy on the time of the entry.

That means that when a user tries to enter a customer record, the MDM Services will prevent any alienation of the Master Data and the Application's record by checking first that the data being inserted is accurate.

2. Information Obfuscation (confusion)
    a. As Sarsfield (2011) continues, sometimes the data errors might not be completely by mistake, but it might be misleading information on the entry form without an alternative option.
    b. The solution to this problem according to Sarsfield (2011) and Herrero et al (2015) is:
        i. **Reward**. "Offer an incentive for those who enter personal data correctly. This should be focused on those who enter data from the outside, like those using Web forms. Employees should not need a reward to do their job. The type of reward will depend upon how important it is to have the correct information" (Sarsfield 2011).
        ii. **Accessibility**. Sarsfield (2011) and Herrero et al (2015) embrace the technologists responsible for the data within an organisation to be open and approachable about criticism from users. If the entry process is not

flexible and understandable by the users, they will find different and incompatible ways to complete their tasks than following the strict rules and complexities.

iii. **Real-time Validation.** Sarsfield (2011) suggests real-time validation as a solution to the problem again.

3. Renegade IT and Spreadmarts (i.e. Spreadsheets that is created and maintained by an individual or group)
    a. What Sarsfield (2011) identifies as renegade IT and spreadmarts are all the independent local applications that some people in the business has approached a person or a group of people from IT outside the IT development roadmap to generate some specific reports because the business is in a hurry to produce some results. These applications do not follow the rules and policies and because of the time constraints are generally not deployed in an orthodox way. This action starts as a quick fix under the surface and usually ends up becoming a "useful" tool over the years. As a result, when the time comes for consolidating the enterprise data, these unorthodox but vital data sources to take time and effort for decoding. The same opinion is shared also with Loshin (2009) who refers to them as Line-of-Business silos, and Dyche (2006) who called them "the new legacy systems".
    b. The solution to this problem according to Sarsfield (2011) is:
        i. **Corporate Culture**. Each action should react, therefore, there should be consequences for those trying to adopt this approach, but there should also be a policy that makes more difficult the creation of these applications and data sources. Also according to Ohbyung et al (2014), the corporation's competence in maintaining data quality ie consistency and completeness positively affects the corporation.

ii. **Communication**. There should be a general education process that enables the employees to understand the negative impact of renegade data.

iii. **Sandbox**. There are cases that the above cause is necessary due to time restrictions and business crucial requirements. However, this process should take place in a controlled and safe environment that is not affecting any production systems; but it is there for experimenting with data, and when the time and the circumstances allow it, to be deployed to a production environment.

iv. **Locking Down the data**. "A culture where creating unsanctioned spreadmarts is shunned is the goal. Some organisations have found success in locking down the data to make it more difficult to export" (Sarsfield 2011).

4. Corporate Mergers

   a. An additional cause that Sarsfield (2011) identifies, is when two or more organisations merge, there is always a case of data quality errors. This process takes place usually very fast, and the IT department who needs to take on this task needs to act immediately without having the required time to plan, the required understanding of the business processes as well as culture clash and differing definitions of the truth between IT departments themselves. Also, this kind of change affect the personnel and has a significant impact, this impact is mostly negative (Stylianou et al. 1996).

   b. The solution to this problem according to Sarsfield (2011) is:

      i. **Corporate Awareness**. "Whenever possible civil division of labour should be mandated by management to avoid culture clashes and data grabs by the power-hungry" (Sarsfield 2011).

ii. **Documentation**. Documentation of the process and data definitions are vital for every IT department to be able to be shared and understood by all the involved members (Sarsfield 2011).

iii. **Third-party consultants**. "Management should be aware that there is extra work to do and that conflicts can arise after a merger. Consultants can provide the continuity needed to get through the transition" (Sarsfield 2011).

iv. **Agile Data Management**. The solutions and the strategies should be kept agile, giving the organisation the ability to be flexible over a merging process (Sarsfield 2011).

5. Corporate Evolution

a. Every organisation's purpose is the continued increase in revenue. This purpose can include the company's expansion, new partnerships and new mergers and acquisitions. Before this actual happens, the organisation has potentially defined what the data quality standards are and at that time, the data quality is defined as fitness for purpose. During the organisational evolution process, this purpose might change. If this process brings new data quality standards or different rules and data integration layers, the fitness for purpose will have to adapt to the new requirements.

b. The solution to this problem according to Sarsfield (2011) is below, however, Mahanti (2018) agrees with Sarsfied in two of the solutions:

i. **Data Governance**. Sarsfield (2011) and Mahanti (2018) agree that by setting up a cross-functional data governance team, the organisation will always have a team who will be looking at the changes the company is undergoing and considering its impact on the information.

ii. **Communication**. As Sarsfield (2011) previously described and Mahanti (2018) agreed, in this instance, documentation is again a very important part to solve this case. As long the documentation is in the right order and

the metadata definitions are well documented, the possibility of bad quality in data should be eliminated.

iii. **Tool flexibility**. Sarsfield (2011) engages the same approach as 4.b.iv as a solution to this challenge.

6. Secret Code
    a. By Secret Code Sarsfield (2011) refers to the actual database content in which in most cases, the application works as it should work while being connected to the local application. However, when integration is taking place between two systems, there might be problems due to shadow data that the database contains. That means that supporting data in one system is not compatible with the other system. In this case, it is difficult to understand all the custom code and special processes that happen beneath the data unless the organisation goes under a profiling exercise.
    b. The solution to this problem according to Sarsfield (2011) is:
        i. **Profile Early and Often.** Don't assume that the data is fit for purpose because it works in the source application. (*In one of the interviews as one of the Business Transformation Director effectively engages to this cause of data quality problems on a very early stage by quoting: "We are not so concerned with taking Data over from ERP1 [This is the Company Two Programme Director from the second audit] if it is inconsistent to start with"*.)
        ii. **Corporate Standards**. Sarsfield (2011) refers to Data Governance Policy following the same example as he has followed previously.
        iii. **Apply reusable Data Quality Tools when possible**. Sarsfield (2011) engages the same approach as 4.b.iv as a solution to this challenge.
7. Transaction Transition

a. By Transaction Transition, Sarsfield (2011) refers to the real-time integrations between systems and especially on the case that by the time that a transaction enters into the database, automatically a process is triggered to send this or these transactions to other downstream databases. However, a malfunctioning system could cause problems with downstream business applications.

b. The solution to this problem according to Sarsfield (2011) is:

i. **Schema Checks.** As part of any Extract, Transform and Load packages, Sarsfield (2011) suggest having a schema checks on both sides of the integrations with controls over the schema validation for both sides. If the schema ensures the validity as per the original plan, then the load should start.

ii. **Real-time Data Monitoring.** Sarsfield (2011) returns to real-time data monitoring as a solution by suggesting frequent data profiling checks as well as database schema checks.

8. Metadata Metamorphosis

a. By Metadata Metamorphosis, Sarsfield (2011) refers to changes that happen between systems regarding the metadata. For example, multiple systems use the same references as product code and product description. If there is a change in a product code in one of the systems, the other systems on the organisation's ecosystem should be aware. (Sarsfield (2011) This is what the research defines as a Unified Data Model where a global referential schema is built on top of all the metadata gathered from all the evolving systems).

b. The solution to this problem according to Sarsfield (2011) is:

i. **Predefined Data Models.** Sarsfield (2011) refers to the industry-specific standards and definitions of what should be in any given set of Data. For example, the automotive industry should follow certain ISO8000 standards, the energy industry should follow the Petroleum Industry Data

Exchange Standards (PIDX) and Office Products industry should follow the BOSS classification or the UNSPSC standards. (BOSS and UNSPSC is the industry standard codebase for office products)

ii. **Agile Data Management.** Sarsfield (2011) comes back to Data Governance as a solution for overcoming this source of the problem.

9. Defining Data Quality
    a. What Sarsfield (2011) means by defining Data Quality is the definition of the data from a holistic point of view across the organisation. For example, sales have a different definition of the product domain within an organisation that the production of the product or the merchandising Loshin (2009) also suggest the same.
    b. The solution to this problem according to Sarsfield (2011) is:
        i. **Standardise Tools.** Sarsfield (2011) and Loshin (2009) suggest using standard tools wherever is possible that these tools aren't tied to a particular solution or application.
        ii. **Data Governance.** Setting up a cross-functional data governance team with people in specific positions to define a common data model.
10. Loss of expertise
    a. Sarsfield (2011) and Maydanchick (2008) refers to specific individuals who have gained the knowledge over the years and understand the meaning of a specific record on a specific field. "Data might be a kind of historical record for an organisation. It might have come from legacy systems. In some cases, the same value in the same field will mean a different thing in different records. Knowledge of these anomalies allows experts to use the data properly" (Sarsfield 2011).
    b. The solution to this problem according to Sarsfield (2011) and Maydanchick (2008) is:

i. **Profile and monitor.** As before the author suggest data profiling for identifying this kind of inaccuracies and monitoring for preventing recurrences.

ii. **Documentation**. Sarfield (2011) suggests that documentation ensures continuity.

iii. **External Consultants**. "Expert employees may be so valuable and busy that there is no time to document the legacy anomalies. Outside consulting firms are usually very good at documenting issues and providing continuity between legacy and new employees" (Sarsfield 2011).

Guess (2011) references Sarsfield's (2011) publication and agrees with all the causes that he articulated. She also agrees with all the solutions that Sarsfield (2011) suggested.

In 2016, Iversen followed the same approach that Maydanchick (2008). He also suggests that there is a categorisation on the way that the data is created just like Maydanchick (2008).

- "Processes that bring data into a database, manually or otherwise, which may either cause problems due to existing, incorrect incoming data or by errors within the extraction and loading processes" (Iversen 2016).
- "Processes that manipulate data already in the database, which can be routine or brought about by upgrades, updates and a range of ad-hoc activities" (Iversen 2016).
- "Processes that cause data to become inaccurate or degrade over time without any physical changes having been made. This usually happens when real-life objects described by the data change while data collection processes remain the same" (Iversen 2016).

However, in his article he identifies as major causes of Enterprise Data Quality Problems only the following seven compared to Maydanchick (2008):

- Initial Data Conversion
- System consolidation
- Batch Feeds
- Real-time Interfaces

- Data Processing
- Data Scrubbing
- Data Purging

Even though Iversen (2016) has described at the beginning the same three-way categorisation as Maydanchick (2008), he focuses only on the two the categories which are “Processes that bring data from outside the databases”, and on “processes that change data from within the database”. The category that he did not analyse is “processes that cause Data Decay”.

McKnight (2009) summarised the reasons that Maydanchick (2008) and Sarsfield (2011) identified the following reasons:

- Entry quality
- Process quality
- Identification quality
- Integration quality
- Integration quality
- Usage quality
- Ageing quality
- Organisational quality

However, McKnight is the first one that refers to MDM directly as a strategic solution to the problem. His strategic approach indicates an MDM Programme as the most effective approach. “MDM provides the framework for identifying quality problems, cleaning the data, and synchronizing it between systems. However, MDM by itself won't resolve all data quality issues” (McKnight 2009). In addition to the MDM Programme, he suggests that “An active data governance programme empowered by chief executives is essential to making the organisational changes necessary to achieve success. The data governance council should set the standards for quality and ensure that the right systems are in place for measurement. Besides, the company should establish incentives for both users and system developers to maintain the standards.” (McKnight 2009). Finalising his article McKnight (2009) suggests that “The result is an

organization where attention to quality and excellence permeate the company. Such an approach to enterprise information quality takes dedication and requires a shift in the organisation's mindset. However, the results are both achievable and profitable." (McKnight 2009).

In 2015 Gulipalli categorised various stages in the data life cycle where deterioration of data quality may occur. These three categories are as follows:

- Data Acquisition
- Internal Data Processing
- Data Manipulation

Gulipalli (2015), follows the categorisation that Maydanchick (2008) initially suggested. He is defining sixteen stages at which data quality problems arise. All these stages follow exactly what Maydanchick (2008) and Sarsfield (2011) described in their publications.

Critical Evaluation

Based on the sources that have been mentioned in this section, all the authors have identified the main reasons for the data quality problems. To summarise, all the challenges arise from the way that data is entering into the system, how this data is getting updated and maintained while in the system and how the data is decommissioned when it is not relevant anymore. It can be mentioned that even if a solution for an existing problem is provided, the organisations and the project teams usually accept the source of the problems, understand the solution but they do not follow the directions that are provided as a solution, even though they agree.

2.3.2. Enterprise Data Quality and Data Governance Methodology

According to Russom's (2006) article 'Taking Data Quality to the Enterprise through Data Governance' which was published at 'The Data Warehouse Institute' (TDWI) on March 2006, one of the most recent methods that companies use to avoid the disruption of their valuable information is MDM. The benefit of this method is that it can cover all types of information that

a company needs such as transactional data, analytical data, and master data. Moreover, companies can add value to their operations when they can manage both transactional and analytical master data as this combination of operations can be the solution of key business problems. The data quality problem that is being questioned can be solved only when the operational side of the business is working together with the analytical side as only this can be the way that the processes can continue functioning without disturbances.

Also, the contemporary architecture of MDM provides tools that eliminate poor data quality such as duplicate identification, real-time change management and data synchronisation. These processes not only minimise the problem of data quality, but they also promote customer relationship management and enterprise resource planning along with business intelligence applications. Regarding the applications that support business intelligence, MDM has the unique capability to create a data model that enables all the data that are attached to applications. It should also be stated (Russom 2006) that in an MDM initiative, data governance is one of the most significant issues, as it incorporates all the management from the availability, utilisation and security of the data in a company. Therefore, from a business point of view, it is considered a priority that data needs to be kept safe but also maintained and exchanged accurately as to fulfil its operational and business cycle. Despite the significance of data to most companies, symptoms of data inconsistency and lack of synchronisation between production systems have been introduced throughout the process and cycle of business operations, business decision making and performance management. In certain occasions, lack of control and management may also result in a violation of regulations.

The control mechanism and systematic approach to maintaining and solving enterprise data problems are known as enterprise data governance. The data governance solution is safeguarded, considering that all requirements and key elements are covered equally.

The first element consists of a ten-process model that together forms a systematic methodology of implementing the solution. The set of processes allows the business to identify, build and share common data definitions, data names and integrity rules. Mapping, analysis, and

profiling of disparate data enable cleansing of data and improvement of data quality. Given all the set of processes and their responsibilities, metadata reflecting those configurations, mappings, cleansings and regulation rules and vocabularies are created. It is worth mentioning that the metadata is accessible to business users to search and share the available defined data. Despite the implementation of processes, the emphasis is given to the management platform that is required for the data governance to be integrated and shared. The technology platform known as the enterprise data management platform consists of a set of tools responsible for managing the data. The last piece of the puzzle is the implementation of data services of the technology platform to deliver a systematic and automated governance process.

Considering the scenario discussed regarding enterprise data governance, a continuous data quality improvement is safeguarded and common processes (as well as re-usable services) are brought upfront to re-enforce the purpose of automated governance.

Griffin (2006) stated that there is an incredible opportunity for organisations that by implementing an MDM solution to standardise their core corporate data across the enterprise; organisations will be in a position to put timely, consistent, accurate and actionable information in the hands of management and knowledge workers. However, one of the challenges that are stated in the article is the concern that the structure that is required to provide ongoing data governance, is too complicated or ineffective. It is a fact that data governance has its costs. However, ineffective governance is much more costly. Other challenges are the definition of the business value of the initiative as well as the reluctance to implement due to the bureaucracy that it creates and the difficulties on prioritising the data standardisation process based on underlying business drivers. (Griffin 2006).

Positive outcomes of data governance are also depicted by the significant growth of sales, cost rationalisation, better working capital management policies and generally greater business processes, overall. Further, the importance of data governance is illustrated by the fact that it is a compulsory requirement following European Union General Data Protection Regulation (GRPR) for companies that have expanded their global footprint. The main factor that forces these companies to employ an MDM initiative which simultaneously requires data governance is the

rising number of mergers and acquisitions. Effective data governance can have positive outcomes and bring solutions when a company operates in different plants in many locations, there are multiple legacy systems or where data exist in multiple languages. Data governance unifies all these various sources of data and facilitates the handling of the information that eventually diminishes the risk of poor data quality.

Critical Evaluation

Based on the sources that have been mentioned on this section, Russom's suggestions are aligned with what Griffin (2006) suggested. Griffin (2006) stated that there is an incredible opportunity for organisations that by implementing an MDM solution to standardise their core corporate data across the enterprise; organisations will be in a position to put timely, consistent, accurate and actionable information in the hands of management and knowledge workers. However, one of the challenges that are stated in the article is the concern that the structure that is required to provide ongoing data governance, is too complicated or ineffective. It is a fact that data governance has its costs. However, ineffective governance is much more costly. Other challenges are the definition of the business value of the initiative as well as the reluctance to implement due to the bureaucracy that it creates and the difficulties on prioritising the data standardisation process based on underlying business drivers. (Griffin 2006).

2.3.3. Implementing Master Data Management

2.3.3.1. Design areas when planning a Master Data Management Function.

Even though MDM is a technical term, the challenges of implementation affect every function within an organisation. Usually, according to Eckerson (2011), the main observation for an MDM implementation is a software implementation even though during this implementation there is a data governance process that affects the overall function of the organisation on a data-consuming basis.

Including the IT element of the MDM process, Joshi (2007), focused on the process of MDM defined eight different that an implementation should involve and follow: These steps are as follows:

1. **Defining the master data flow** (Joshi 2007). Data owners should define the source and destination systems that contain master data.
2. **Identifying the generation and the consumers of master data** (Joshi 2007). The applications that generate master data and the end consumers of this data, should be identified.
3. **Collecting business metadata** (Joshi 2007). The core entities and their core attributes including any data types or constraints and dependencies should be captured.
4. **Defining the master data model** (Joshi 2007). That includes the current form of the master data, how this data is going to be mapped and how it will be established on the master data model.
5. **Defining the functional and operational characteristics of the tool** (Joshi 2007). Based on the organisation's requirements the functionality of the Master Data Management tool should match the business functions.
6. **Merging the source data to create a master data element** (Joshi 2007). This integration process requires a significant effort from the business to validate that the source data is matching the master data after the integration and transformation following the business rules that have been applied.
7. **Collection and maintaining metadata between technical and business rules** (Joshi 2007). The testing process should be initiated immediately after the transformation, ensuring that the generation of master data, does not create any errors during the generation process.
8. **Publication of the Master Data** (Joshi 2007). The master data is now ready to be consumed by any function that requires access within the business.

Based on the above eight steps (Joshi 2007) Vilmico (2013) and Pekkola (2013) defined a ten-step approach for a Master Data Management implementation. These steps are as follows:

1. **Identifying the need and objectives** (Vilmico and Pekkola 2013). The first step is to understand the needs that lead to a Master Data Management implementation. Based on these needs, the business should understand the changes within their functions that such an implementation would cause to the business. However, the emphasis should be on the cost savings that would be achieved by:
    - More effective work.
    - Improved reporting.
    - Service-oriented architecture interoperability.
2. **Identifying the organisation's core data and processes that use it** (Vilmico and Pekkola 2013). This step is divided into four different parts.
    - Part 1. The definition of master data and what is the difference between master data and transactional data.
    - Part 2. The establishment of specific criteria for master data and the discovery of the master data sets within the organisation.
    - Part 3. Analysis of existing information systems and analysis of their data. This part leads to application mapping. These are the applications that hold the core data and execute the main processes within the organisation.
    - Part 4. Observation of the services and the processes that are associated with the data.
3. **Defining governance** (Vilmico and Pekkola 2013). According to Vilmico and Pekkola (2013), governance can be defined in three levels. Each level had its roles and responsibilities.
    - The organisational level. This level requires a concept owner who is the lead in developing the Master Data Management. The concept owner in most cases is also the sponsor for the project. Within the organisational level, Vilmico and Pekkola (2013) also identified the need for an operative role who will be responsible for leading the development and the implementation of the Master Data Management process.

- Support function level. This level focuses more about the privacy and the security of the data as well as the ongoing maintenance of the quality control within the information systems and the integrations.
- Data set level. This level which has been identified as essential is related to the business units that consume the master data and are involved in the Master Data Management process and development.

4. **Defining the maintenance process** (Vilmico and Pekkola 2013). "Master Data Management processes refer to the processes that are needed for administrating and maintaining master data. This includes the responsibilities, methods and tools for collecting data (e.g. forms), defining workflows and guidelines for reviewing data in the workflows, and appropriate instructions for users and administrators" (Vilmico and Pekkola 2013).
5. **Defining the data standards** (Vilmico and Pekkola 2013). Data standards should be applied in the content and the model of master data set on an attribute level. The data model is the enabler for applying any changes within the business process. Moreover, the model is consumed by the applications that consume this data as well as the reporting within the organisation.
6. **Metrics for Master Data Management** (Vilmico and Pekkola 2013). Data quality is the main metric for a Master Data Management implementation as it describes how well the data serves the organisation's demands. Data qualities definitions should be formed to generate specific policies that should be enforced to master data sets. Data cleansing and data migration strategies should be defined in this step to ensure the continuity of the Master Data Management process.
7. **Planning a Master Data Management architecture** (Vilmico and Pekkola 2013). Master Data Management architecture should contain information about all the applications that are involved. Information about data flows between the systems and data administration practices and points. The Master Data Management architecture should address any security and privacy issues. There are three models for a Master Data Management

architecture these are the a. repository, b. registry and c. hybrid approach. In the next subchapter 3.4.2. "Master Data Management architecture" Obenhofer and Dreibelbis (2008) agree with Viswanathan (2006) that the three models for Master Data Management architecture are registry, coexistence and transactional which is different from what (Vilmico and Pekkola (2013) defined.

8. **Planning training and communication** (Vilmico and Pekkola 2013). Communication with everybody involved is important for a successful MDM implementation. Communications and training plans should be included within the schedule and a strategy of how the Master Data Management and data governance will be communicated should be important.
9. **Forming a roadmap for Master Data Management development** (Vilmico and Pekkola 2013). This step is focused on the development phase of the Master Data Management implementation.
10. **Defining Master Data Management applications' functional and operational characteristics** (Vilmico and Pekkola 2013). Back to step 7 however, the following components should be considered as functional and operational characteristics. These components are user interface, workflow, and Master Data Management functionalities such as data creation, editing, removal and reconciliation.

2.3.3.2. Master Data Management Architecture

Although data governance as an MDM initiative facilitates the transactions and offers a more convenient use of information with limitation of poor quality data, many issues should be handled with attention as this is a complicated process (Viswanathan 2006). The implementation of an MDM is a process that requires the harmonisation of three different factors of each multinational enterprise: a) the people which are the users and the people responsible for the implementation, b) the operations which are composed by the requirements and the

specification of the main operations of the business and c) the extent of the international presence of each company that uses different sources of data.

IBM was one of the pioneers to provide master data solutions and has shown a keen interest to embark for new opportunities. In cooperation with TATA PLC consultancy, IBM published the article 'A Methodology for Sustainable Success with Master Data Management Initiatives ' (Viswanathan 2006).

Viswanathan (2006) talks about the conceptualisation that Service Oriented Architecture (SOA) and MDM share common ideas. This statement is based on the cohesion of SOA to decompose an application into a collection of reusable services after going through stages of organisation education, exploration, and evaluation of available technologies. After the investment in SOA Pilots to validate the chosen technologies, businesses stumble onto the fact that the “Master Data” is scattered around multiple applications. This is where MDM is initiated, therefore becoming a milestone in the stages of SOA deployment.

The importance of this deployment is for the business to better understand the management of master data across different contexts (Viswanathan 2006). While there is a range of initiatives responsible for their problems, a solution suggests that those problems are combined as one, forming the idea of “commonly shared concern”. It makes perfect sense that the above initiatives are opposing to those of an Application Implementation where the master data has been acquired and managed from multiple applications.

Under these circumstances, a new set of concerns are brought to the attention involving data governance, data quality, data model and the data lifecycle that is de-coupled from the domain of the individual applications. Undoubtedly data governance can easily be discussed further as an urge to face dilemmas that are often controlled the business.

As far as the application is concerned, a radical transformation will be subject for transformation due to the de-coupling of the master data from the application. In the same way, implying that the master data is no longer part of the application but a service (Master Data Service) that is available to be accessed externally.

Based on the above article, Obenhofer and Dreibelbis (2008) defined three different methods of use for the MDM as well as three different Implementation styles which taxonomise and answers the concerns that Viswanathan (2006) had in his article. These methods of use by Obenhofer and Dreibelbis (2008) are the following:

- **Collaborative method of use:** "Collaboration means that multiple users, usually in different roles, participate in the same process on a master data entity." "Key requirements of a collaborative method of use are workflow support with check-in/check-out functions, support for relationships, and product hierarchy management. From a security perspective, the attribute-level granularity of authorisation privileges across all functions such as workflow, relationship and hierarchy management must be available for implementation" (Obenhofer and Dreibelbis 2008).
- **The operational method of use:** "This method of use is important when an MDM System has to function as an Online Transaction Processing (OLTP) server. Typically, a large number of applications and users require quick access to master data to retrieve and change master data through MDM services invoked by business processes" (Obenhofer and Dreibelbis 2008).
- **Analytical method of use:** This method is divided into three different sub-methods:
    - **Identity analytics**: "This sub-type is usually encountered when there is a need to determine or verify identity and discover hidden relationships" (Obenhofer and Dreibelbis 2008).
    - **Analytics on Master Data**: "Here, an MDM System needs to answer questions such as 'How many new customers did I receive over the last day?' or 'How many customers changed their address in the last week?'" (Obenhofer and Dreibelbis 2008).
    - **Analytics integration with Data warehouses**: This is a combination of both the above sub-methods which combines the referential integrity and the analysis on the Master Data within a Data Warehouse (Obenhofer and Dreibelbis 2008).

Based on the methods of use that have been described above there are also three different implementation styles which answer the concerns that Viswanathan (2006) mentioned in his article. These Implementation styles by Obenhofer and Dreibelbis (2008) are the following:

- **Registry Style**: In which the Master Data Management solution identifies and captures the inconsistencies of the data and reports any bad quality data.
- **Coexistence Style**: in which the MDM identifies and captures the inconsistencies of the data and through the Master Data Management solution the users can push these inconsistencies to the relevant systems.
- **Transactional Style**: in which the MDM solution is integrating with the applications in real-time and can apply the MDM policy on the time of the entry. That means that when a user tries to enter a customer record, the MDM services will prevent any alienation of the master data and the application's record by checking first that the data that is going to be inserted is accurate.

Critical Evaluation

Based on the sources that have been mentioned in this section, all the authors supplement the suggestions that each one of them makes. In addition, Viswanathan (2006) raises concerns during the project implementation and Obenhofer and Dreibelbis (2008) define the three different approaches based on the nature of the project that is going to be implemented. Both approaches have been used in the past during an MDM implementation. The overall outcome, however, is that the suggestions from Joshi (2007), Vilmico (2013) and Pekkola (2013) combined with Viswanathan (2006) and Obenhofer and Dreibelbis (2008) are aligned with the MDM Framework that has been explained in detail in section 2.2 of this chapter. All the six steps that are defined within the Framework are supported by what the authors suggest regarding the design and the planning of the MDM process as well as the main architecture based on the method of use and implementation style. Both are key decisions that will define the entire strategy and roadmap on how the process and the programme will have to be planned and scheduled.

2.3.4. Why Data Quality matters

Every business decision for an IT project must be taken after a detailed investigation on technology, platforms, and professional contractors and of course a purpose why this project needs to take place and what the Return on Investment is. In 2011 Talend published a white paper identifying the 10 reasons that Data Quality matters and can help businesses to achieve greatness at their projects. The paper covered the advantages of improved data quality across the areas of the enterprise including improved business intelligence, enhanced data governance, risk management analysis and compliance to expanded sales as well as more efficient supply chains and lower operational costs.

In this white paper, Talend (2011) identifies as the 10 most important reasons as follows:

1. **Data profiling and understanding at the beginning of the projects**. Bearing this in mind, it will be a great advantage on each implementation to have already performed a cleansing, standardising and de-duplication exercise with the data at an early stage as this could cause severe delays on the project or even completely jeopardy (Dorr, Murname 2011). There is an instance on the presentation of the results that describes the danger of not focusing on the data quality at the beginning
2. **The privilege to own high-quality data** will give to the project a better and more accurate view of all the data in the data repository or on the data silos that each enterprise owns. Especially for the data warehouses and the business intelligence solutions, high quality of the data will give a standardised set of functions that make dissimilar data sources shareable giving a higher level of confidence in the business forecasts knowing that the results are based on truthful data (Dorr and Murname 2011).
3. **Ability to build better customer relationships and create future upselling opportunities based on an accurate, de-duplicated customer dataset**. The management of quality on attributes like name, address, contact telephone numbers, email addresses and the other customer-related attributes that defining a single version of the customer can make every

Customer Relationship Management system more effective by using accurate information (Loshin 2011).

4. Based on the previous reason. Talend (2011) takes a step further and analyses **the extra costs that the enterprise may obtain based on customer data** containing typos misfiled data or wrong addresses. These costs can occur when it comes to shipping products, market to the customer list, sending recall notices or cross-selling and upselling the customer base. Also, shipments that contain incorrect shipping information may cause penalties from the shipping vendor or even the customer itself (Dorr and Murname 2011).
5. **Every IT project involves data migration.** Even if it is a completely new system or an upgraded version of a current system. On every project, the first thing that needs to be completed is the project plan. In the project plan, the data profiling task exists to provide and complete understanding of the nature of the data before the scheduled task of migration. Data profiling enables the project team to understand the difficulties of the migration task and prompt the team to create a better and more accurate plan (Dyche and Levy 2006).
6. Going back to Customer Relationship Management, Talend (2011) states that the importance of credit decisions can be to the customers and how these decisions can affect the **customers' satisfaction and happiness.** Based on the previous sentence, Talend (2011) refers to the amount of data that drives credit decisions. These decisions are driven by customer credit quality measurements, agency debt ratings account receivables and, or current market exposure. When the data is integrated with multiple systems which have different data standards, different formats and naming conventions, the unification of the data on individuals and businesses who may ask customers for credit can delay or even worse can make a wrong decision. High quality of data can help the business and each team to identify, validate and manage the quality of data that are taken under consideration for all credit decisions (Lohr 2011).
7. The **accuracy in reporting data** for customers and their sales figures. It is known that every business must deal with regulations and compliances from government agencies and

regulatory bodies that they set up numerous strict regulations. Violations of these regulations might result in lawsuits, government action or bad publicity that damages the organisation's reputation and credibility. High-quality data is the key to managing this risk by ensuring the proper controls to prevent compliance is in place (Folmer et al, 2014).

8. The **"One Big Picture"**. What Talend (2011) means by that is that there are many occasions that companies are facing difficulties in reviewing and understanding their performance. For example, a comparison between two financial reports that are supposed to show the same numbers may result in different figures. In most cases, data quality is usually the reason. By engaging on high-quality rules and by profiling and discovering data anomalies, structure and overall suitability before any data migration begins, it needs to be ensured that the data across multiple systems and domains are matched and the relationships have been identified. By applying the above, the enterprises know that the metrics for their key performance indicators are based on healthy data feeds and data quality metrics (Dorr and Murname 2011).
9. **The supply-chain efficiency**. Data quality is a necessity for an enterprise resource planning system consolidation due to the reason that the data in the system presents accurately or inconsistently the inventory levels. The organisation's inventory carrying costs depend on the level of accuracy of the data as well as its health. If the data on the organisation's inventory data is wrong, the business cannot support real-time delivery and of course real-time levels of stock. At the same time by ensuring high data quality on the organisation's inventory, the organisation can apply improved intelligence about the business and corporate buying power.
10. **Data governance**. By data governance, an organisation ensures that everyone within this organisation is striving to deliver more accurate information. Teams and departments within the organisation proactively manage data and understand its value. This is an exercise between business and IT and how these two separate areas, can cooperate and work in partnership to achieve a resolution on the problems of data management (Koltay 2016).

Overall, this publication states why the quality of the data is important in every industry and every function of a business. The total cost of data inconsistencies is far more expensive in the long run than resolving these inconsistencies at the beginning.

Following the above, House of Brick Technologies (2014) summarised in three different categories the main reasons these are the following:

- **High Cost of Poor Data Quality Data**. As an example, the article analyses the UPS Corporation case study and focuses on the incorrect addresses and mailing costs. Based on the article, there is an annual increase in the mailing costs even for bulk mailing rates. The article states that bulk mail costs increased 6% only in January of 2014. Taking into consideration this annual increase, UPS published that 23% of all their mailing addresses were incorrect and in conjunction with the increases the quality of the data is of the highest priority. The costs of the organisation are massive. Except for the costs on the external operational level, the article identifies that the internal costs presented that approximately 80% of the code that was used on the systems, was written to handle data anomalies. The article continues on the bases that if the data that is coming in the program is correct, the development and time cost can be reduced as well as the program maintenance bringing more flexibility in the application's creation and making IT clients more productive.
- **Data quality's Impact on business**. The article focuses on customer satisfaction, but it is expanding on the way that customer retention and potential loss of sales are measured. If a customer had an unpleasant experience due to incorrect data, this customer will pass the feedback to other colleagues or friends, and these people to other people. Effective up-selling, handling refunds, government regulations, and risk reduction are just a few of the other areas being impacted by poor data quality. "Correcting data quality in any size company is a daunting task. Without senior management support, it will fail. Implementing data quality changes, how a company views data. It is the largest asset most companies have, and some have gone so far as to include it in their balance sheets." (House of Brick Technologies, 2014)

- **Data Quality is a lifetime task**. The article describes the data quality task as a never-ending effort. “Even though many people have tried, you cannot boil the ocean. Data quality cannot be solved all at once. Attempting to do so will almost always result in failure and discouragement.” (House of Brick Technologies, 2014). (On chapter 3.5 Sherman R identifies as “Boil the Ocean” as one of the worst practices)

Eckerson (2009), as the Director of Education and Research for the Data Warehousing Institute (TDWI), suggested that “…poor quality data can have a deleterious impact on the health of a company. If not identified and corrected early on, defective data can contaminate all downstream systems and information assets, jacking up costs, jeopardizing customer relationships, and causing imprecise forecasts and poor decisions” (Eckerson 2009). The problem is that data changes over time. It is suggested that over two per cent of customer data are becoming outdated monthly due to the changes in the circumstances of the customers. “More perniciously, as organizations fragment into different divisions and units, interpretations of data elements mutate to meet the local business needs. A data element that one individual finds valuable may be nonsense to an individual in a different group” (Eckerson 2009).

Eckerson (2009) states in his article that:

> *“[P]art of the problem is that most organisations overestimate the quality of their data and underestimate the impact that errors and inconsistencies can have on their bottom line. On one hand, almost half of companies believe the quality of their data is ‘excellent’ or ‘good.’*

Yet, almost half of respondents on the interviews said the quality of their data is "worse than everyone thinks.

Critical Evaluation

Based on the sources that have been mentioned in this section, all the authors have supplemented each other on the importance of Data Quality. As it is stated in the section, the impact of bad data quality within an organisation can be catastrophic. The cost of poor data quality can affect multiple areas within an organisation not only in relations to the way that a company interacts with their customers and suppliers but on the way that internal operations

are deadlocked due to data inaccuracies. Internal misuse of good quality data can lead to increased cost on Capital and Operational expenditure. This expenditure will have to be spent on overcoming the internal operations and to an extent any external operations. That is why the investment in ensuring the data quality is necessary and ongoing since the maintaining of data quality must be a never-ending task.

2.3.5. What needs avoiding: Worst practices in Enterprise Data Governance.

Moving forward from the top 10 reasons of why data quality is important, and taking into consideration the last reason, reason 10 with regards to data governance, Sherman (2011) (founder of Athena IT Solutions) with more than 20 years of IT experience and speaker in industry events, published an article with the worst practices in enterprise data governance and a guide of what must be avoided.

Starting his article, Sherman (2011) emphasises on the fact that organisations have started realising that the data that is stored is getting accumulatively increased, but the analysis and the insight of this data are not possible for most of the organisations. The problem that the Sherman (2011) sees is the transformation of all this data into meaningful information that meets the 4C i.e. comprehensive, consistent, correct, and current information. Sherman (2011) clarifies that the above problem is not a problem that technology can solve for the organisation, but that technology can be the starting point for an organisation, to establish a data governance programme that the data can be truly treated as a corporate asset, by enforcing consistent definitions, rules, policies and procedures.

Sherman (2011) proceeds by clarifying that this is the goal for each organisation and that many companies have launched many efforts to apply an enterprise data governance policy, but the results of these attempts were not successful and encouraging. The reasons that he identifies as pitfalls or worst practices are below considering them as “red flags” to alert that a data governance implementation might be leading in the wrong path.

- **“Buy-in but not a commitment”.** As the first and most important worst practice point, Sherman (2011) identifies the “buy-in” from the business. The business accepts and understands the reason of why a data governance programme needs to be done, they put the people in the right places but when it comes to them to do their tasks, the engagement is very limited to none. People from the business side need to create data definitions, to create the business rules and of course the key performance indicators for the data governance programme. Once they define and agree on the above, the business people need to enforce their usage and compliance and ensure that the definitions, rules and KPIs are updated on an ongoing basis as the business evolves and changes. In most of the cases, the reality is that the data governance tasks are assigned to business managers with an already overloaded schedule as a result, these tasks to be placed lower on the level of importance that they should be. “Without a real business-resource commitment, data governance will take a back seat to the daily firefight and will never be implemented effectively” (Sherman 2011).
- **“Ready, Fire, aim”.** As the second point on the list, the Sherman (2011) identifies a correct approach most of the times in the beginning by the business people and recognises that process of creating a governance steering committee. Also, separate governance working group is the appropriate task. Usually, the steering committee is appointed by business representatives from across the business and the working group is usually made up of the data stewards who are the body that performs the real governance labour. Sherman (2011) identifies timing, as an important point that organisations usually do wrong. As he identifies, the organisations are usually mistaken on trying to assign the people and form these panels before the in-depth understanding of what the scope and the real purpose of the data governance programme are, and the roles and responsibilities of the participants. “A guaranteed way to stall a data governance initiative in its tracks and lead the business to lose interest is to prematurely organize the management framework and then realize you need a do-over” (Sherman 2011).

- **"Trying to solve world hunger or boil the ocean".** On the third point of his worst practice list, Sherman (2011) identifies one very common problem that exists in nearly all organisations. Sherman (2011) identifies as a trap on the most data governance efforts the results that are expected by the business people. Most of the business bodies by starting a data governance problem they believe that it is feasible to solve all of an organisation's data problems in the initial phase of the project. Or that the organisations will start with their biggest data problems, issues that distance the entire enterprise and are likely to be very political. Sherman (2011) states that it is almost impossible to establish a data governance programme while at the same time the organisation tries to solve data problems that have taken years to build up. In this case "you need to think globally and act locally" (Sherman 2011). "In other words, data problems need to be broken down into incremental deliverables. "Too big, too fast" is a sure recipe for disaster" (Sherman 2011).
- **"The Goldilocks Syndrome".** "In the story of Goldilocks and the three bears, the little girl keeps encountering things that are either one extreme or another, which is precisely what happens on many data governance programmes" (Sherman 2011). Thus in most cases, either the foundations of the programme are too high level and substantive data issues are never really dealt with, or usually the organisation's bodies are trying to create data definitions and regulations for every data field in every table in every application that an enterprise has. As a result, the programme is dragged down by so many details that it ends up being impossible to complete. In these cases, there is a need to be a common agreement between the bodies, and to achieve a compromise between these two extreme situations that will make the data governance programme an asset to the business.
- **"Committee overload".** Sherman (2011) in this point recognises the fact that there are many people involved in the programme from all the aspects of the business. Usually, though, this large number of people turns to be over-crowded. Usually, based on statistics, he states that the highest volume of people involved, the chance for more

politics and more watered-down governance responsibilities becomes higher. For a successful implementation, Sherman (2011) recommends a limit of people between 6 and 12 and these people there are bodies with the required decision-making authority.

- **"Failure to implement".** Sherman (2011) takes a step further and indicates the most common thing to happen after the data definitions, the business rules and the KPIs have been created. This is usually the non-enforcement of the above in any business process. In this case, the data governance effort will not produce any business value. As per Sherman (2011), the data governance process should be recurring feedback, in which data is defined, monitored, acted upon and changed when appropriate. Also, the other problem at this point is that after the setup of the above definitions and rules, there is no communication about the governance initiatives. This situation can easily lead to old habits by the business users and the data governance programme losing momentum.
- **"Not dealing with change management".** At this point, Sherman (2011) addresses another aspect of the current processes of both IT and Business. He identifies as a risk the lack of need for change management procedures. Again, he articulates the internal politics as the main problem on the change management process and that these challenges need to be overcome.
- **"Assuming that technology alone is the answer".** Sherman (2011) at this point analyses the wrong approach that most businesses have regarding the solution of a problem with the purchase of a Master Data Management, data integration or data quality software (or even a mix of the three), as a solution to eliminating the data governance problem. The combination of high price tags of the vendors' software usually sets high expectations with the most expected outcome to be the avoidance of "nasty people, processes and political issues" (Sherman 2011). He clarifies that an organisation may find value on purchasing a high-end software but, in reality, it is the internal processes and interactions that make or break the effort of a data governance programme.
- **"Not building sustainable and ongoing processes".** Sherman (2011) in this point, emphasises on the fact that the businesses most of the time are willing to proceed with

the initial investment in time, money, and people. However, this investment most of the time is for that time and with no future assigned budget or resource commitments or even processes for the future.

- **"Ignoring Data Shadow systems".** As a final point, Sherman (2011) refers to the "shadow" systems that can make the difference. One of the most common mistakes in a data governance programme, is the focus and emphasis on the "Enterprise Transactional Systems" and the Business intelligence systems, assuming that every other system are not important and most significantly all the data can be found "there". Often though, key information is located in data shadow systems that exist within the organisation.

In addition to Sherman's ten worst practices in Data Governance (2011), Woodie identifies five mistakes that all of them match with Sherman's (2011). These are the following:

- "**No Data Governance Strategy**". Woodie (2016) suggests that "Data governance refers to an overarching strategy that defines how organizations ensure the data they use is clean, accurate, usable, and secure" (Woodie 2016). He describes an approach that most of the business follows, which is the process of solving the data issues on an ad-hoc basis when a problem appears. He defines the components of a Data Governance Strategy as:
  - "setting up processes that dictate how data is stored and protected" (Woodie 2016)
  - "setting up a set of standards and procedures for ensuring how authorized personnel can access and use data" (Woodie 2016).
  - "and setting up controls and procedures to ensure the rules are being followed" (Woodie 2016).

He concludes his first point by highlighting that data governance does not work without continuous improvement and he suggests that for a successful strategy, the organisation needs to start in a smaller scale and grow it over time.

His first point matches with three of Sherman's (2011) points and these are the ready, fire and then aim approach that goes under the "solving issues on an ad-hoc basis". The second one is the "Not building sustainable and ongoing processes" which goes under the fact that data

governance does not work without continuous improvement and the third is the "trying to solve world hunger or boil the ocean**"** which goes under the fact that each organisation needs to start a data governance programme in a smaller scale and grow it over time.

- "**Relying too much on Unicorns**". Woodie (2016) suggests that most of the companies turn to data scientists for everything that has to do with data, expecting them to turn raw data into actionable insights overnight. He suggests that "Data governance is best led by a collection of data stakeholders from the IT department, line of business, and compliance. The Data Governance Institute also recommends hiring a Data Governance Officer (DGO)" (Woodie 2016). Woodie (2016) matches with the first point from Sherman (2011) "Buy-in but not committed". He identifies that a data governance programme is a team effort for both IT and business and he highlights that "Without a real business-resource commitment, data governance will take a back seat to the daily firefight and will never be implemented effectively" (Sherman 2011).
- "**Letting Schemas Run wild".** Woodie (2016) suggests that the mistake usually happens on the implementation of the data repositories. The "schema on reading" approach is very important to data governance principles that require knowing what kind of data is stored and being processed. This matches with Sherman's (2016) bullet point of "Goldilocks syndrome" which indicates that it can either be too much detail or too less detail. It also matches with the "Failure to implement" approach because the data governance process, should be completely recurring feedback in which data is defined, monitored, acted upon and changed when appropriate. When these standards are not met, then there is an increased potential for failure. The above bullet point can be associated with the "Not dealing with change management**"** since there lacks detailed planning from the business or the project team.
- **"Storing Everything forever".** Woodie (2016) suggests that one of the most important parts of a data governance strategy is data retirement. Organisations very often suggest that they want to "Keep Everything" resulting in a huge spend of their budget on storing data that they will never use. This matches again with Sherman's (2011) points of "Goldilocks

syndrome" and "Trying to solve world hunger or boil the ocean" which indicates too much or too little detail.

- **"Not using Power Tools".** Woodie (2016) identifies that the right people need to be in the right position to implement an effective data governance strategy. A good and effective policy, that indicates the priorities and the processes that will help in implementing the data governance is also required. However, an extra step is the selection of the right tool. Just like Sherman (2011) with his "Assuming that technology alone is the answer", Woodie (2016) identifies that no tool will solve every data governance challenge on a day by day basis. However, specific tools can help automate a substantial part of it.

Following Sherman (2011) and Woodie (2016), Bajkov (2016) identifies thee major worst practices that are associated with Sherman's (2011) and Woodie's (2016).

- **Technology**. Data governance and data management is not only a software but is a 3-P process (People, Politics and Process). In this statement, Bajkov (2016) combined all the above bullet points.
- **Speed Trap**. "Many organizations want fast results and want to make sure that they can snag the best people for the project. They assign all of the roles and responsibilities to staff before having established a project scope or developed an understanding of what is appropriate. A data governance project is very public, and the project loses support every time a change is required" (Bajkov 2016).
- **All or Nothing**. Bajkov (2016) in a very simple statement suggests that if the plan is to solve every data problem in the initial phase, the project will fail. "The process needs to be global in planning and incremental in phasing" (Bajkov 2016).

Critical Evaluation

Based on the sources that have been mentioned in this section, all the authors supplement the suggestions that each one of them makes. From the sources that have been explored in this

section, the worst practices can be all grouped and categorised in a three-level hierarchy. The **first group** can be the 'Speed Trap' as previously defined, which can be associated with the Relying too much on Unicorns and no Data Governance Strategy. Both of them can also be expanded to not building sustainable and ongoing processes, not dealing with change management, failing to implement a solution and "Ready, Fire, Aim" as described through the section. The **second group** can be the 'All or Nothing'. The sub-categories that can be associated with this group is the uncontrolled schemas, which can also lead to storing any kind of data and finally as per the first group the relying too much on Unicorns. Both of them can be expanded and related to the previously described worst practices "trying to solve world hunger or boil the ocean", "Goldilocks syndrome", "Buy-in but not commitment" and "Committee overload". The **third group** can be 'Technology'. The sub-categories on this group can be associated with not using any power tools and as the previous groups with the relying too much on Unicorns. The worst practices that can be linked to this group is the "failure to implement", the "ignoring the data shadow systems" and the "assumption that the technology alone is the answer" to all the challenges.

Category	Sub Category	Worst Practice
Speed Trap	Relying too much on Unicorns	Not building sustainable and ongoing processes
	No Data Governance Strategy	Not dealing with change management
		Failure to implement
		Ready, Fire, aim
All or Nothing	Letting Schemas Run wild	Trying to solve world hunger or boil the ocean
	Relying too much on Unicorns	Committee overload
	Storing Everything forever	Buy-in but not a commitment
		The Goldilocks Syndrome
Technology	Relying too much on Unicorns	Ignoring Data Shadow systems
	Not using Power Tools	Assuming that technology alone is the answer
		Failure to implement

Table 4. Worst practices in Enterprised Data Management – Critical Evaluation.

The above grouping indicates that the three categories are connected. In summary, all these problematic practices indicate that there is always a risk when Senior Management and programme teams assume that technology can resolve all the organisations' challenges in a short period. Technology can provide the solution to the challenges when is combined with the

appropriate resources that understand the process and a programme with a plan of actions that indicate no risk to one another in the right time.

2.3.6. The political landscape of MDM

As it has been stated previously, every BTP has many difficult stages. Since there were multiple technicalities on the previous sections of this chapter this section is focused on how to overcome **politics and political problems based on the attitude of each business function against the new approaches**. These new approaches bring changes and in every business, every change brings emotion and an emotional reaction (Ladley 2012). Whenever a business activity crosses an organisational functional area or division, there will be political issues, and all organisations are political (Smith 2008). Conflict is inevitable, but good politics allows for the healthy resolution of conflict and the continued development of a functional organisation.

Based on the above, Smith (2008) highlights the following eight points to better understand the political landscape within an organisation.

- **What is the current organisational structure**? Know the levels on the organisational chart and how each interacts (collaboratively, authoritatively, decentralized, etc.). Is the organisation fragmented and is each portion operating independently for the capture and use of data and information? How do the leaders at each level interact (up, down, across)? What are their attitudes toward each other and how are those attitudes affecting the organisation? How are decisions made, communicated and enforced (Smith 2008)?
- **Is there social synergy in the organisation**? Are things arranged so that individuals and groups add to (rather than subtract from) the long-term evolutionary potential of the organization (Smith 2008)?
- How can the organisation be structured to support and maintain the charter, sense and values of data governance (Smith 2008)?
- **How can the communities in the organisation increase collective intelligence**? Reflect on the historical collection and use of data; problems and opportunities to be explored

for improved data definition and cross-organizational usage, and develop methods to learn from its experiences to create appropriate activities and relationships with data and information throughout the organisation (Smith 2008).

- **How can the organisation's governance system maintain an evolving, coherent approach to data management**, rather than drifting into social incoherence (misnamed "anarchy") or some rigid and dysfunctional status quo (Smith 2008)?
- **What social feedback systems are missing**, blocked or dysfunctional which, if present in healthy forms, would allow the organisation to regulate data management consistently (Smith 2008)? (Examples of social feedback systems: success criteria, economic indicators, accurate information about organizational conditions, institutionalized collective self-reflection, acquisition of externally generated perspectives, etc.)
- **Is the governance system capable of generating and revising shared vision** and culture as the need arises? If not, why not and what can be done to correct this "bad political climate" (Smith 2008)?
- **What is the quantity and quality of organisational dialogue**? Are all relevant viewpoints and stakeholders involved and is dialogue conducted in the spirit of contribution (Smith 2008)? (Dialogue is defined as multi-directional communication to enhance shared understanding — as contrasted with debate [communication to win], and all forms of one-way communication.)

Focusing on MDM and data governance, the communication between all the functions that are involved is very important. "Data governance programmes especially focus on authority and accountability for the management of data as a valued organisational asset. Data governance should not be about command and control, yet at times could become invasive or threatening to the work, people and culture of an organisation" (Seiner 2014).

As part of the research and as part of the audits and the interview that will follow, the research will also focus on how the members of an organisation react to the introduction of this new authoritative process.

During the audits and the interviews (chapter 4), most of the people involved were trying to avoid any element of data governance or MDM; or they were trying to reduce the element of responsibility that belonged to them. There was a misunderstanding of what data governance requires and how it can be implemented. According to Seiner (2014), the data governance should make clear that its purpose is to help the people of the organisations do their jobs but also helping them data stewards understanding that they were not being given any responsibility beyond what they already had.

The three main concerns that organisations have based on Seiner (2014) are as follows:

- "Most organisations view data governance as something over and above normal work efforts that threaten the existing work culture of an organisation" (Seiner 2014).
- "Most organisations have a difficult time getting people to adopt data governance best practices because of a common belief that data governance is about command and control.
- "Data governance is the execution and enforcement of authority over the management of data. Nowhere in this definition states that data governance has to be invasive or threatening to the work, people and culture of an organisation" (Seiner 2014).

Following the above, Ladley (2012) suggested five activities to overcome the problem of emotional reaction to change during a Data Governance initiative. These five activities are as follows:

- **Get the right sponsor in place**. The right sponsor will be able to facilitate communication more efficiently (Ladley 2012).
- **Appreciate the frequent and open communication**. Commonly, companies will try to withhold information out of fear of employee reactions based on the idea that if people do not know activity will not react and as a result will be easier to accomplish. However, this approach will create rumours and organisational churn which will make it even more difficult to control in the future (Ladley 2012).

- **Appreciate that there is a psychology to change**. Understanding how people react is essential to structuring the Data Governance initiative to deal with it (Ladley 2012).
- **Be clear and specific about what is changing**. Based on this, it will be easier to understand what is required to be done to overcome behavioural changes. People cannot change behaviour if they are unaware of what they are supposed to change (Ladley 2012).
- **Realign performance objectives and accountabilities**. The change will not be completed successfully unless people are held accountable for required behaviour changes.

Critical Evaluation

Seiner (2014), Smith (2008) and Ladney (2012) all agree that politics, communication and a common vision are few of the most important attributes to successfully achieve Data Governance. There are three critical components to the political view of data governance: objective high-quality analysis, publicly visible forums for the management of the data governance and stewardship functions, and organisational engagement about data issues. Each contributes to the application of "good" politics to influence the acceptance and practice of data governance, and the absence of any of these components can cause the failure of a governance effort (Smith 2008). This demise will cause emotion and the emotion will cause a reaction. The reaction, in this case, will be the programme to be sabotaged at every turn by those who are opposing it or see the programme as a threat (Ladney 2012).

Multiple conversations were formed during the interviews and the audits. As part of these conversations, three steps have been followed to understand how people were reacting to MDM and data governance as part of the BTP but also what their perspective was on the subject. These steps are as follows:

- The first step was to educate the people in the business areas on what MDM and data governance is and how this process will benefit their organisation and their daily tasks.

- The second step was to ask specific questions that prompted them to speak about things that they could not do and any issues that they were facing regarding data that they define, produce and use.
- The third step was to document the outcomes of these conversations and demonstrate back to them how what they discussed is going to improve the business and add value to their process.

This approach of non-invasive data governance that was defined by Seiner (2014), has been applied to an extent during the experimental data collection. Concluding this section it is to be highlighted that, "Politics is the art of the possible" (Smith 2012) and data governance needs healthy politics to achieve its mission of managing the availability, usability, integrity, and security of the organisation's data.

2.3.7. Concluding Summary

This Chapter introduced the foundations of MDM, the reasons why organisations should see the MDM process as a necessity and focused on the reasons that the MDM initiative is necessary.

There were formal definitions for the reasons behind the data quality problems and there were multiple representations of opinions in which all of them had complimented each other. Also, there was an analysis of the Enterprise Data Quality and Data Governance methodology and a demonstration that effective data governance can have positive outcomes and bring solutions when a company operates in multiple different environments.

There was also a demonstration of the MDM architecture and design areas of focus when planning the MDM. Also, there was an analysis of the importance of Data Quality and an investigation and comparison of opinions regarding the pitfalls that need extra careful prior, during and after the implementation of data governance.

The conclusion to the chapter was the political landscape within the organisation and how the transition to a more structured set of rules can be adopted in a less to non-invasive way to the broader business and end-users.

2.4. Key findings between MDM Framework and Theoretical Foundations.

At the beginning of chapter 2.2. it has been stated that the proposed framework is based on Infotech Research Group (2014). The proposed framework has been used in the past in multiple instances and has been proven successful. In this section, there is going to be a discussion between the proposed framework and the theoretical foundations of MDM.

The proposed framework which is based on figure 1 (pg. 12) and is explained in detail in chapter 2.2. is divided into 6 different phases. These phases are as follows:

1. Making the case for MDM
2. Preparations for MDM
3. Understanding the MDM needs of the business
4. Define and maintain the MDM architecture
5. Create an MDM strategic roadmap
6. Maintain the MDM process

MDM is not an IT project. It is a process that affects all the functions of the organisations, that is why the first part of the process is to establish the foundations of why this process is necessary. The importance of MDM, as well as the benefits and the responsibilities for both business and IT, have to be defined and be very clear (Longman 2008). This statement is also supported by Eckerson (2011) when he suggested that the main observation for an MDM implementation is a software implementation. However, during this implementation, there is a data governance process that affects the overall function of the organisation on a data-consuming basis.

The second step on the framework is the steps that need to be taken to prepare for the MDM. None of the authors in chapter 2.3. suggested what Infotech Research Group (2014) suggested as a second step. One of the most important parts of the second phase is the maturity

assessment that is described in section 2.2.2.3 is required for determining the readiness of the MDM. This maturity assessment is used to identify potential areas within the organisation that need development. Based on the results of this assessment, the needs can be identified in more detail.

Based on the above, the third phase that is defined within the framework is the understanding of the master data needs within the business. This step is aligned with Vilmico and Pekkola (2013) and their first step. In their 10 steps approach the first step is to understand the needs that lead to an MDM implementation. Based on these needs, the business should understand the changes within their functions that implementation like that would cause to the business. Also, the third phase of the framework is aligned with the third step from Vilmico and Pekkola (2013) which is identifying the organisation's core data and processes that use it. In this phase of the framework, a clear definition of data domains is required as well as the differentiation between master data and reference data.

The next step from Vilmico and Pekkola (2013) is the definition of data governance. The proposed framework suggests that data governance should be defined in a later stage. The reason for that is that a clear understanding of the architecture should be in place before any governance (Weber et al 2009). The data governance affects people's behaviour and based on Seiner (2014) there are three main concerns.

These concerns are as follows:

- "Most organisations view data governance as something over and above normal work efforts that threaten the existing work culture of an organisation" (Seiner 2014).
- "Most organisations have a difficult time getting people to adopt data governance best practices because of a common belief that data governance is about command and control" (Seiner 2014).
- "Data governance is the execution and enforcement of authority over the management of data. Nowhere in this definition states that data governance has to be invasive or threatening to the work, people and culture of an organisation" (Seiner 2014).

If the architecture is not in place with a very clear definition of what is required and what processes are going to be affected, there is a risk that the data governance would be invasive and the reaction from the end-users could be unpredictable. That is why is believed, that in this instance the framework provides a more secure option on the definition of the data governance. For that reason, the fourth step of the framework suggests that the next phase should be the definition of the architecture.

The fourth phase of the framework aligns with Joshi's (2007) eight steps process that an MDM implementation should involve. However, all the steps that Joshi (2007) suggests separated between the fourth and the fifth phase of the framework which is the creation of a strategic roadmap. These two phases are also aligned with Vilmico and Pekkola's (2013) seventh step, which is the planning of an MDM architecture. However, the ninth step from Vilmico and Pekkola (2013) which is the formation of a roadmap is also aligned with the fifth phase from the proposed framework.

The last part of the framework is also aligned with the last step from Vilmico and Pekkola (2013). The framework suggests as the last phase the maintenance of the MDM. Vilmico and Pekkola (2013) suggest on their last step the maintenance of functional and operational characteristics.

Overall, the proposed framework which originally got introduced by Infotech Research Group (2014), is supported by the literature on its technicalities. However, there is an element that is not included in the literature. This element is the human reaction in change. The literature suggests the steps that need to be followed. However, it is not taking into consideration, how invasive or not their approaches can be. Every technicality that is suggested within the literature exists within the framework. The difference, however, is that the framework has been designed taking into consideration the way that any change can be imposed on the different functions within the business.

Framework	Theory	Alignment
MDM is a process that affects all the functions of the organisations	During the MDM implementation, there is a data governance process that affects the overall function of the organisation	The importance of MDM, as well as the benefits and the responsibilities for both business and IT, have to be defined and be very clear.
Prepare for MDM	Determine Readiness	This maturity assessment is used to identify potential areas within the organisation that need development. Based on the results of this assessment, the needs can be identified in more detail.
Understanding of the master data needs within the business	Identifying the need and objectives.	Based on these needs, the business should understand the changes within their functions that implementation like that would cause to the business
	Identifying the organisation's core data and processes that use it	A clear definition of data domains is required as well as the differentiation between master data and reference data
	Defining governance	A clear understanding of the architecture should be in place before any governance.
Define and Maintain Data Integration Architecture	• Defining the data standards • Metrics for Master Data Management • Planning a Master Data Management architecture • Planning training and communication	If the architecture is not in place with a very clear definition of what is required and what processes are going to be affected, there is a risk that the data governance would be invasive and the reaction from the end-users could be unpredictable.
Define and Maintain Data Integration Architecture	• Defining the master data flow. • Identifying the generation and the consumers of master data. • Collecting business metadata. • Defining the master data model. • Defining the functional and operational characteristics of the tool. • Merging the source data to create a master data element. • Collection and maintaining metadata between technical and business rules. • Publication of the Master Data.	This integration process requires a significant effort from the business to validate that the source data is matching the master data after the integration and transformation following the business rules that have been applied

Create an MDM Strategic Roadmap	• Forming a roadmap for Master Data Management development. • Defining Master Data Management applications' functional and operational characteristics	Once the implementation has been completed from a development point of view, strategic roadmap documentation needs to be created based on the solution that has been delivered. This roadmap will define the use of the system, maintenance of the system and general management that will be defined on the next and last stage of the MDM implementation lifecycle
Maintain MDM	Defining the maintenance process	This includes the responsibilities, methods and tools for collecting data (e.g. forms), defining workflows and guidelines for reviewing data in the workflows, and appropriate instructions for users and administrators
Not in Framework	The human reaction in change.	• Most organisations view data governance as something over and above normal work efforts that threaten the existing work culture of an organisation. • Most organisations have a difficult time getting people to adopt data governance best practices because of a common belief that data governance is about command and control. • Data governance is the execution and enforcement of authority over the management of data. Nowhere in this definition states that data governance has to be invasive or threatening to the work, people and culture of an organisation.

Table 5. Key Findings between MM Framework and Theoretical Foundations

Chapter 3: Research Methodology and Implementation of the Research

3.1. Introduction

This chapter aims to provide insight into how the research was designed and the environments that the experimental data collection took place.

As the main part of this research, a set of twenty-eight semi-structured interview questions were devised and thirteen hours of interviews with senior management were recorded. The research focused on the analysis from the results and tried to present the challenges when it comes to the implementations of the MDM and data governance.

Based on the results of the interviews and with comparison to the literature review and analysis, the research tried to identify if there is a problem on the flow. If there is one, how this affects the process.

In addition to the interviews, there were two data audits. These audits have been performed as an initial assessment to understand the level of maturity with regards to data strategy within the business in different stages of a BTP. Also, these audits were performed to identify the level of data quality and to establish a foundation on how each business is using its data. Moreover, to understand the level of governance that each organisation follows in gathering, processing, storing, and using its data.

3.2. Research Methodology

This is an overt deductive research project using qualitative methods.

3.2.1. The Research Philosophy

There are many research philosophies available that can be applied in this study. Investigating in more detail the current context, epistemology is covering three key philosophies: positivism, realism, and interpretivism.

Positivism philosophy will not be used in this research as this philosophy is focused on researches that have to do with natural sciences (Saunders et al 2007).

The definition of realism philosophy states what the senses understand as reality is the truth. The objects exist autonomous from human perception. The fundamental approach of realism is that reality is independent of what the human mind understands. Realism assumes a scientific approach to the development of knowledge (Saunders et al 2007).

Realism can be separated into two different types, direct realism and critical realism. The first approach which is direct realism states what it is perceived, and what can be seen. The second approach which is the critical realism focuses on what it can be felt with the senses, for example, an image of an object in the real world but not the entities directly (Saunders et al 2007).

Interpretivism philosophy involves researchers to interpret elements of the study, that is why interpretivism integrates human interest into a study. Accordingly, “interpretive researchers assume that access to reality (given or socially constructed) is only through social constructions such as language, consciousness, shared meanings, and instruments” (Myers 2008). This philosophy emphasises qualitative analysis over quantitative analysis.

Interpretivism is “associated with the philosophical position of idealism, and is used to group diverse approaches, including social constructivism, phenomenology and hermeneutics; approaches that reject the objectivist view that meaning resides within the world independently of consciousness” (Collins 2010).

According to the interpretive approach, the researcher as a social actor needs to appreciate differences between people. Moreover, interpretivism studies usually focus on meaning and may employ multiple methods to reflect different aspects of the issue (Dudovskiy 2018).

Interpretivist approach is based on the following beliefs:

- **"Relativist ontology.** This approach perceives reality as intersubjective that is based on meanings and understandings on social and experiential levels" (Dudovskiy 2018).
- **"Transactional or subjectivist epistemology.** According to this approach, people cannot be separated from the knowledge; therefore, there is a clear link between the researcher and the research subject" (Dudovskiy 2018).

With the Interpretivist approach, qualitative research areas such as cross-cultural differences in organisations, issues of ethics, leadership and analysis of factors impacting leadership can be studied at a great level of depth. Primary data generated via Interpretivism studies might be associated with a high level of validity because data in such studies tend to be trustworthy and honest (Dudovskiy 2018).

Based on all the above, the research philosophy that will be adopted in this research is the interpretivism.

3.2.2. The Research Approach

According to Saunders (2012), there are three different research approaches, the deductive, the inductive and the abductive approach.

Moreover, Saunders (2012) suggest that:

- If the research starts with a data collection for exploring a phenomenon to generate or build a new theory, then the most appropriate approach is the inductive (Saunders 2012).
- If the research starts with a data collection for exploring a phenomenon, identifying themes and explain patterns for generating a new or modifying an existing theory which

subsequently there is a testing and additional data collection, then the most appropriate approach is the abductive (Saunders 2012).

- If the research starts with a theory developed from reading the academic literature and there is a research strategy designed to test this theory, then the most appropriate approach is the deductive (Saunders 2012).

The following table summarises the three different approaches from Saunders (2012).

	Deduction	**Induction**	**Abduction**
Logic	In deductive inference, when the premises are true, the conclusion must also be true	In inductive inference, known premises are used to generate untested conclusions	In an abductive inference, known premises are used to generate testable conclusions
Generalisability	Generalising from the general to the specific	Generalising from the specific to the general	Generalising from the interactions between the specific and the general
Use of data	Data collection is used to evaluate propositions or hypotheses related to an existing theory	Data collection is used to explore a phenomenon, identify themes and patterns and create a conceptual framework	Data collection is used to explore a phenomenon, identify themes and patterns, locate these in a conceptual framework and test this through subsequent data collection and so forth
Theory	Theory falsification or verification	Theory generation and building	Theory generation or modification; incorporating existing theory where appropriate, to build a new theory or modify existing theory

Table 6 Deduction, induction and abduction: from reason to research (Saunders 2012)

The inductive research approach will have to be discarded because the research does not start with a data collection to generate a new theory. The abductive research approach will also have to be discarded because there is not a subsequent data collection after the initial data collection.

The approach that is most relevant to this research is the deductive research approach. This approach starts with existing theory and the research design is structured based on theories in the research field and approaches from recognised sources and institutes (for example TDWI,

DAMA, DGI). In this document, there will be an examination of what MDM and data governance is, and how the field professionals approach this process. How these professionals approach the MDM methodology and how the business and commercial aware employees perceive the definition of MDM. How this conflict between these two groups generates more challenges and how the results of these miscommunications drive to the stability of the successful implementations of MDM solutions to less than a quarter (24%) of the recorded global MDM implementations (Gartner Research 2019). According to Saunders (2007), all approaches can be used but in different stages.

During the research lifecycle, currently, available literature will be used to evaluate the current theory and proceed to data. An evaluation of the relationships between the most common MDM approaches that are used will be tried to be achieved. Mainly qualitative data will be used to answer the research objectives by critically evaluating the data and finally to come to conclusions.

Bryan and Bell (2003) stated that to ensure the successful process of case study and research approach there are three basic criteria: validity, reliability, and replicability. Following this, controls will be applied to guarantee the validity of the data that will be used to reach a conclusion that will be completed after the primary and secondary research. There will be an attempt to warrant the clarity of the definitions and will examine the specific outcome of the investigation by proving the theory or by indicating the necessity for modification.

During this research, two different data audits have been conducted in two different organisations involving multiple people across different positions. The audits used as an enabler to act as a social actor that appreciates differences between people and offered access to reality through social constructions such as language, consciousness, shared meanings, and instruments.

In addition to the audits, this research includes twenty-eight semi constructed interview questions that have been asked to a different group of people during the different project. These people were:

Roles: There were eighteen participants with diverse backgrounds and coming from a different function of the business. The roles of the interviewees are divided in:

- 2 different Chief Information Officers (IT Directors),
- 3 different Chief Financial Officers (Finance Directors)
- 1 Master Data Management Director,
- 3 Programme Directors for 3 different Business Transformation Programmes
- 2 IT Project Managers
- 1 Head of Business Analysis.
- 4 Business Analysts.

Project Types and Scales: Due to the nature of the Organisations that the interviewees were coming from, the monetary scale of these projects was summing up to millions of pounds. However, the focus was on projects that were part of a BTP. The average years of experience of each interviewee are 12.75 years which is distributed on an average of 20 years per CIO and CFO, the MDM Director has 23 years of experience, and the Heads of Development are on an average of 17 years. Project Directors have an average of 8 years of experience. These 8 years are only the years that they have as Project directors excluding any prior Project management roles. The same approach applies to the IT Project Managers which from the average of 6 years of experience any experience before their current role is excluding from counting. The Head of Business Analysis has 4 years of experience as Head of the function and 13 years as a Business Analyst. Also, the 4 different Business Analysts that got interviewed had an average of 4 years of experience per analyst.

Based on the above roles, the qualitative research areas such as:

- cross-cultural differences in organisations,
- issues of ethics,
- leadership and analysis of factors impacting leadership could be studied.

Moreover, due to the role within each business of the interviewees, the data that has been collected can be characterised trustworthy and honest.

3.2.3. The Primary and Secondary Research

This research identifies the primary research method as very essential and very appropriate with regards to the answers that are expected. According to Aaker (2004), primary data is defined as the collection of this kind of information that is used to answer a specific research question and match the research objectives. There are multiple different methods for collecting primary data that can be used to examine a theme or a hypothesis like qualitative research, surveys and tests. As an addition, based on experience, it has been identified that a different technique for collecting data could lead to false outcomes, so according to the previous statement, the method that matches better the current research is a case study.

To achieve a comparison between the currently available literature with the approaches that currently have been used to implement an MDM solution, there was a need to investigate case studies that refer to MDM implementations (Data Audits). Also, there was a need to schedule several personal interviews with business and commercial people as well as IT Professionals (like database professionals and information administrators). At the same time, communication to businesses that there was any involvement in the past by achieving a successful or failing implementation of MDM and data governance.

Except for the Primary research, this research will apply a secondary research method which is fundamental for the outcome. Current arguments and perspectives that are stated in the literature had to be investigated thoroughly to achieve the research aim.

The resources that were used include written materials like reports, surveys' results, public records, books, journals, magazine articles. This information can be accessed by material that is owned as well as materials that Sheffield Hallam University Library holds. For accessing the required Journals, Sheffield Hallam University offers the Sheffield Hallam University Literature Search Online scheme. Also, for case studies and past implementations of MDM, notes from his previous involvements can be used as well as case studies from TDWI, the Search Master Data Management Research, and Gartner Research. The books will be used for the theoretical aspects of the research and the matter-of-fact examples will be retrieved from more up to date information available on the web to support the theoretical part of the research.

3.2.4. The Research Choice

In summary, the research has been conducted based on two audits, multiple semi-structured interviews and very limited use of a questionnaire which is used as an enabler for more conversation.

As previously stated in the primary and secondary research definition, this type of data collection has been used to approach the required information to answer the research objectives and the request questions.

The structure of the research (as it has been identified up to present) directs the research to apply mono method technique such as qualitative methods. Quantitative research method analyses and investigates existing theories after the examination of statistical and numerical data to empower the currently available literature of MDM and data governance (Saunders et al. 2007).

Examining the nature of the current research, the research method that is most appropriate to the current study is qualitative research. During this research, qualitative data was collected and processed.

The investigation has consisted of semi-structured interviews with field professionals and business employees. The information that was gained was qualitative.

Strictly structured questions decided not to be included because this approach can cause limitations during the interview process as a result not to gain the information that is required.

There are parts in this research that could be characterised as action research. These parts are focused on the audits that have been performed. **The research consists of two anonymous audits**. The main goal of both audits was to define and understand the current state of the MDM and data governance state of the organisations during and after the implementation of a BTP. The following steps during the audits have been followed:

- **Plan**. Initially, time spent with the sponsor of the audits to understand the requirements and plan the body of the investigation

- **Act**. Actions had been taken as planned by interviewing the relevant people that were involved in the audit process and explained to them the reason that these audits took place
- **Observe**. Observations and recording of the responses from the interviews had to be occurred to produce the audit reports for submission to the audit sponsors.
- **Reflect**. The audits concluded by presenting the results to the sponsors and critically reflected in the outcomes.

Even though there are some elements of action research based on the four steps described above. The role was changing after the initial circle that is why this research cannot be characterised as action research.

3.2.5. The Research Feasibility

This research is likely to be delivered successfully. The reasoning behind this statement is that the research consists of two already approved data audits focused on MDM and data governance process during BTPs. In addition to the audits, there have been twenty-eight semi-structured interview questions that have been performed to eighteen industry-specific people in multiple roles. The results from the interviews and the data that has been collected can be characterised trustworthy and honest due to the role within each business of the interviewees.

3.2.6. The Research Ethics

As part of the data audits, both organisations agreed to the use of these audits for the research purpose. However, they requested that any reference to their business or employees should be completely anonymised. In this research, the two audits that are used as part of the research have been completely anonymised. Company One and Company Two have been used to refer to the companies that the audits took place.

As part of the semi-structured interviews, every participant was made aware of the purpose of the interview and that the entire conversation was being recorded and that the entire

conversations would be a transcript for this research. Even though the recordings include names and roles of each participant, none of the participants' names has been used. On the presentation of the results, each person is referenced by their role.

Please refer to appendix 6 for the Ethics Checklist.

3.3. The Research Strategy

The current research is separated into two different phases. The phase of examination and the phase of testing. During the examination phase, there will be a study on the MDM approaches and the implementation methodologies, and the most common decisions taken by the IT professionals and the business and commercial aware users. How these decisions are made and why there is no common ground between these two objects. Also, there will be an effort to understand the difficulties that most of the projects face and why it is so expensive and difficult to implement an MDM solution.

In 2007 Saunders stated that the most important step for any research is to choose the most appropriate strategy for its topic. However, this statement does not indicate that the research strategy is more important than others. In practice, it is fundamental to approach a study by a research strategy that can answer the research question and the research objectives of the research. The research question and objectives including the time frames will be the reason for the selected research strategy. These factors are the guides to choose the strategy.

Case study or case studies in this research is the strategy that has been selected for his investigation during the testing phase. The case studies in this research will be the data audits and the interview answers since they were made during a BTP. The generation of answering to questions like why what, how, who, when what and where are attributes of the case study research strategy. This ability defines the case study research strategy as a strategy which involves an empirical investigation of a contemporary occurrence within its real-life outline using several sources of facts.

3.3.1. Research Positioning

A BTP consists of three distinct phases. The first phase is the selection of the appropriate set of application systems as well as the introduction of new business processes across multiple lines of business and different channels. The second phase is the implementation of the new application systems and the data migration process. The third and final phase is the transition from the legacy application systems and business processes to the newly defined framework of processes and technologies that ensure the business and data continuity. MDM encompasses a pivotal role during the second and the third phase of a BTP and is defined as the process that runs in parallel with any other business process; assigning responsibility to people and technology on processing, capturing, maintaining and defining data accuracy based on a defined set of rules.

As it is stated in section 4.1. this research performed experimental data collected from a set of twenty-eight semi-structured interview questions and two data audits that are used as case studies. These audits have been performed as an initial assessment to understand the level of maturity with regards to data strategy within the business in different stages of a BTP.

The First audit is on phase three of a BTP after a non-successful implementation. The research through the audit investigates on the points of what went wrong and how these points can be addressed on that stage.

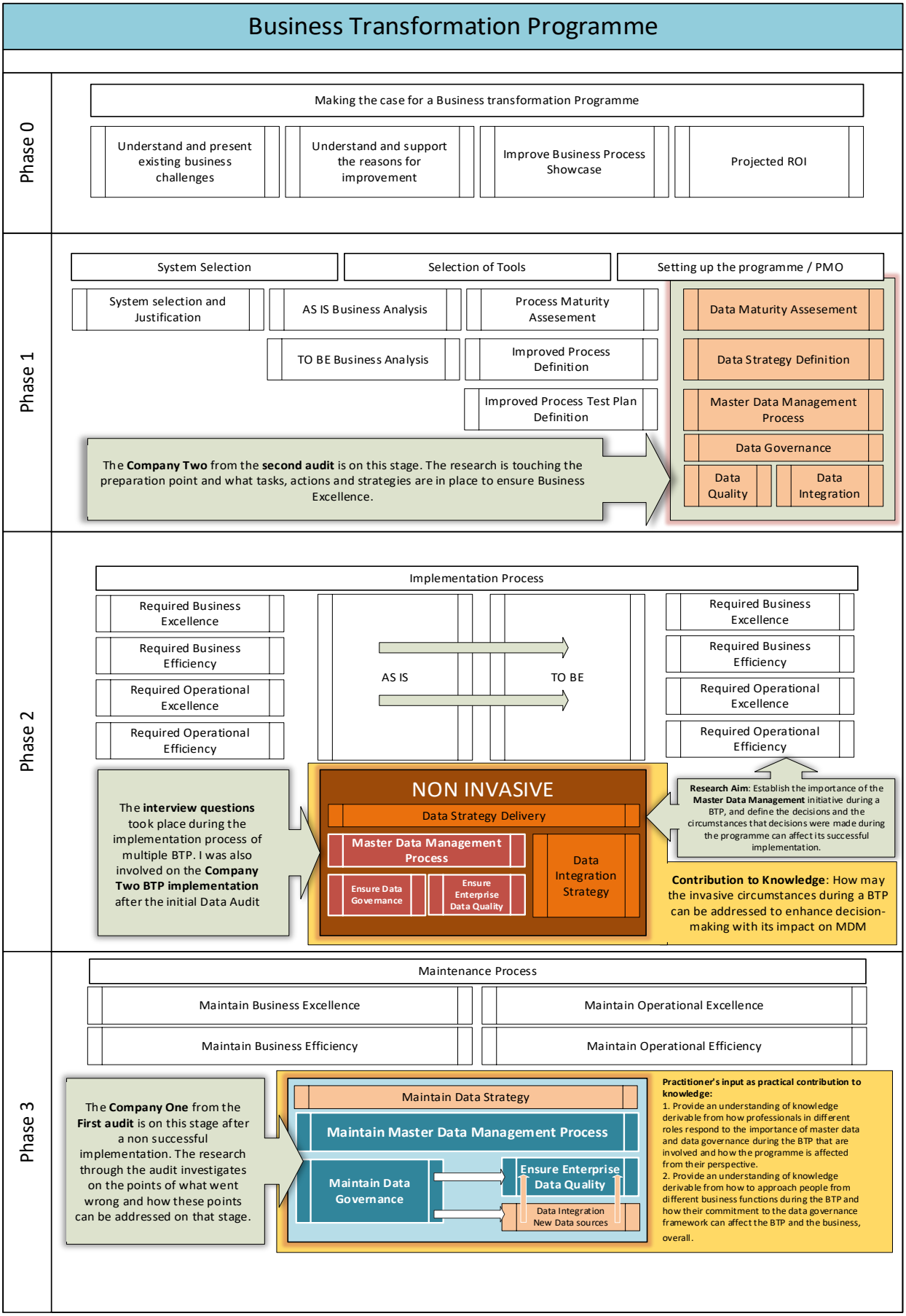
The second audit is in phase one. The research is touching the preparation point and what tasks, actions and strategies are in place to ensure Business Excellence.

The interview questions took place during the implementation process which is phase two of multiple BTP.

The following diagram highlights the areas that the research took place. Even though a BTP takes place within three phases as described above, there is a fourth phase which takes place before. This phase is Phase 0 on the diagram and includes all the initial conversations within a business and ensures the motivation and the reason that a transformation is required. None of the interviews or the data audits took place in phase 0.

These four phases that are described on the diagram, have been identified after multiple involvements and interactions with BTPs in experimental level.

Figure 2 Research positioning on the next page.


Business Transformation Programme
Phase 0
Making the case for a Business transformation Programme
Understand and present existing business challenges
Understand and support the reasons for improvement
Improve Business Process Showcase
Projected ROI
Phase 1
System Selection
Selection of Tools
Setting up the programme / PMO
System selection and Justification
AS IS Business Analysis
Process Maturity Assesement
TO BE Business Analysis
Improved Process Definition
Improved Process Test Plan Definition
Data Maturity Assesement
Data Strategy Definition
Master Data Management Process
Data Governance
Data Quality
Data Integration
The Company Two from the second audit is on this stage. The research is touching the preparation point and what tasks, actions and strategies are in place to ensure Business Excellence.
Phase 2
Implementation Process
Required Business Excellence
Required Business Efficiency
Required Operational Excellence
Required Operational Efficiency
AS IS
TO BE
Required Business Excellence
Required Business Efficiency
Required Operational Excellence
Required Operational Efficiency
The interview questions took place during the implementation process of multiple BTP. I was also involved on the Company Two BTP implementation after the initial Data Audit
NON INVASIVE
Data Strategy Delivery
Master Data Management Process
Ensure Data Governance
Ensure Enterprise Data Quality
Data Integration Strategy
Research Aim: Establish the importance of the Master Data Management initiative during a BTP, and define the decisions and the circumstances that decisions were made during the programme can affect its successful implementation.
Contribution to Knowledge: How may the invasive circumstances during a BTP can be addressed to enhance decision-making with its impact on MDM
Phase 3
Maintenance Process
Maintain Business Excellence
Maintain Operational Excellence
Maintain Business Efficiency
Maintain Operational Efficiency
The Company One from the First audit is on this stage after a non successful implementation. The research through the audit investigates on the points of what went wrong and how these points can be addressed on that stage.
Maintain Data Strategy
Maintain Master Data Management Process
Maintain Data Governance
Ensure Enterprise Data Quality
Data Integration New Data sources
Practitioner's input as practical contribution to knowledge:
1. Provide an understanding of knowledge derivable from how professionals in different roles respond to the importance of master data and data governance during the BTP that are involved and how the programme is affected from their perspective.
2. Provide an understanding of knowledge derivable from how to approach people from different business functions during the BTP and how their commitment to the data governance framework can affect the BTP and the business, overall.

3.3.2. Design of Audit Reports

During this research, there were two organisational audit reports. The reason behind these two audits reports was that one organisation was in the middle of a business transformation programme, and the second organisation was at the beginning of a business transformation programme. Both organisations requested an audit on the MDM and data governance process and how this process would integrate with the BTP of each business.

The audits have been conducted to solve specific problems that had appeared during the programmes within both organisations. In the first company, the BTP was facing challenges due to data inconsistencies and false decisions. In the second company, the business transformation programme needed an external consultant to review the current process and plan the data strategy for the business moving forward.

The methodology that was used for each audit was based on the "Data Audit Framework" by Jones, Ross and Ruusalepp (2009). "The data audit framework enables organisations to find out what data they hold, where it is located and who is (or is not) responsible for it, as well as offering a tool for managing this information and sharing it with other organisations in a controlled environment" (Jones, Ross and Ruusalepp 2009).

For each Audit Report, the specific structure was followed to complete a report.

The audit approach was developed in five steps as follows:

1. **Step 1.** Each audit started with an introduction explaining the reason that this audit report is performed followed by an executive summary of what has been found. The executive summary is always required depending on the level of seniority of the audience. Usually, when the audience is part of the company's board of directors of a C level executive, the attention had to be taken at the beginning since this type of audience has limited time to perform tasks outside their already set schedule. That is why the executive summary needs to be simple, complete and be able to point only what the receiver is interested in.

2. **Step 2.** After the executive summary, a reference to the background of the company was needed to be presented to justify and establish the main reason of why an audit is required as well as to ensure that the background has been fully understood by the auditor.
3. **Step 3.** After the presentation of the background, the next step that needed to be presented was the aims and the objectives of the audit report. The aims and the objectives were defined in the initial interview with the person who requested this audit (this person is usually a Director or C level executive). Based on the initial interview the auditor had to log the desired requests /desired outcomes of the person who asked this audit and based on these requests the auditor had to set up the architecture of the Audit and the method that is required.
4. **Step 4**. Once the aims and objectives of the audit have been defined the auditor moves into explaining the method that he is going to follow during this audit based on the requirements that the sponsor defined.
    a. The first stage of each audit was to undertake a preliminary risk assessment which included Interviews with the company's management.
    b. The second stage of each audit was the planning stage, which included multiple interviews with key employees that currently run the MDM and data governance process. The conversations were based on the aims and objectives that came out from the first interview within the preliminary risk assessment stage.
    c. The third stage of each audit was the Testing phase which is the review of the interviews and the processes.
    d. The fourth and final stage of each audit was the representation and the wrap-up phase of the audit methodology.
5. **Step 5.** After the description of the methodology that was followed during the audit, the next step was to present the results and the conclusions as well as suggested improvements and next steps back to the main sponsor.

3.3.3. Design of Semi-Structured Interview questions

The second part of the research was based on interviews. Out of the twenty-eight questions, twenty-five were used as an enabler for a discussion. The remaining three had the format of a Likert scale questionnaire, **however**, this Likert scale questions have been used again, as an enabler for more discussion and **not as a quantitative** data gathering tool. The main purpose of the questions was to understand and measure the level of understanding within the business and within the respective projects regarding data quality and master data. The participants consisted of eighteen people with different managerial and technical roles as well as project management roles.

At this section of the Thesis, each question is followed by an explanation of the question and a reason of why these questions have been asked. Moreover, in this section, there is a presentation of each question that has been discussed with the participants. For each discussion point, there is an explanation of what the research is trying to achieve with the question as well as the main aim of each question.

During the interviews, each participant was made aware that these questions are going to be used for this research, and they made aware that they had been recorded. Additional notes except the recordings have been taken during the interviews. The five questions that had the format of a questionnaire on a Likert scale, have been also recorded. Even though these three questions were based on a Likert scale, there was still a conversation point between the participants.

At the beginning of each interview, the interview document was being distributed to the participants to assist with the questioning. (The document can be found in appendix 5)

Starting from question one to question twenty-two, were interview questions, which were part of the semi-structured interviews. The rest of the questions (twenty-three to twenty-eight), were part of a questionnaire which was used as an enabler to enrich the quality of the discussion. Question twenty-four is including in the questionnaire, however, it does not represent a Likert scale format. The question twenty-four was a supplementary question to twenty-three and all the participants answered in a very simple way. The Aim, definition, and reasoning for each of

the twenty-eight questions on the questionnaire are below. The statistics presented on some questions have been taken by Gartner Research (2008).

1. ***What is your Role***

The role of each interviewee had to be recorded for the interviewer to understand and classify the role that each interviewee was going to play on the project.

2. ***What kind of projects have you worked on? What scale?***

The scale of the projects and the type of projects had to be clear to understand the expectations and the exposure that the interviewee had with projects regarding data quality and master data.

3. ***What is the importance of Data Quality on these projects?***

Based on the answers that the interviewee would give on the second question the interviewer had to understand the perspective of the persons regarding data quality and the importance of their project in combination with their role. This would give a better view of what the interviewees believe is better for their project with regards to process, data, or both.

4. ***Are you aware of Master Data?***

A question about the Data Quality was asked before the interviewees' understanding of master data. The reason for that was to understand if they separate master data and data quality or if their main focus is on transactional data and they are interested more in the quality of the process and not the actual data the process requires.

5. ***What is your definition of Master Data?***

Based on the answers to all the previous questions. There was a need to understand if the interviewee is aware of the master data as a term and if they can differentiate the master from the transactional/operational data.

6. ***What is your view of Master Data Management and how important is it for your Project?***

Bearing in mind that all the interviewees were involved in implementing highly data-driven systems as the main initiative of their projects. There was a need to understand how big of a part the master data was taking into their projects and how important for them was the "Data Take On" strategy before the initiation of the project.

7. ***Are you aware of the 3 classifications that a business can be regarding MDM?***
    - *Best in Class*
    - *Industry Average*
    - *Laggard*

In this question, there was a need to identify and, in some cases, to start a discussion/debate on how the MDM process can be measured and classified. This classification could be achieved based on the time that it takes to establish the rules for processing quality data, how they identify inconsistencies, how do they deal with duplicates, inconsistent data and inaccurate information and of course how much does it cost to perform all this process.

8. ***Are you aware of the KPIs with Regards to MDM?***
    a. *From 1 to 10 (1 low 10 high) what is the mark that you would give to your business with regards to the following?*
        i. *Formal MDM system in Place*
        ii. *Automated Capture and Creation of Data*
        iii. *Cross-functional team both IT and business lines to guide Master Data Management Implementation*

In this question, there was a need to understand mainly how the interviewees were measuring the quality of their data to achieve the objectives of their highly data-driven projects and then how they set up any key performance indicators with regards to MDM. This was mainly a question to understand how the interviewees establish measurements on their data to evaluate their

process. The way that they add measurements on their process indicates how they measure the quality of their transactional data and as a result, how they measure the quality of their static data that forms the transactional data.

9. ***How would you describe your data strategy with regards to specific business objectives?***

MDM is a key element that involves People, Process and Technology. That is why a strategy/plan to bind these three elements together is required. On a highly data-driven system implementation such as Enterprise Resource Planning, Warehouse Management System or Customer Relationship Management system the data is the key part. There was a need to realise how the interviewees consider the data as part of their overall process strategy, project strategy or if they have a different data strategy that they follow and out of that strategy the project and process strategy is defined.

10. ***How would you describe the support from the senior management with regards to the MDM improvement?***

Following the conversation from the ninth question, MDM is a combination of People, Process and Technology and on any large organisation, the people play the most important part. People make decisions, people set up strategies and people follow processes to achieve a business objective; that is why the support from the decision-makers is important for any project initiative. There was a need to understand the relationship between the decision-makers and the people responsible for delivering the projects as well as the technical people and how the suggestions from technical people are absorbed by non-technical people and the opposite.

11. ***How would you describe the tools that you implemented to manage and maintain Data Consistency?***

Based on the turn that the interview would have taken after ten questions, there was a need to investigate into a more technical ground. This question was focusing more on the process within the MDM process. This question was targeting into understanding the next steps after the completion of the project and how the "after project lifecycle" of data integration would operate.

12. ***How would you describe the Cost, the time and the complexity of an MDM implementation and what is the ROI of High-quality data?***

With this question, there was a need to understand how the interviewees understand the MDM and the data quality in general. This question was mainly focused on the perspective that people have for data against the business process and if people are more interested in investing in the process instead of understanding that the value of the information that the process generates depends on the quality of the data that is involved.

13. ***How much time do you think that an employee of your organisation will spend on searching for data per hour/week?***
    a. *Best in Class is 1.2 hours /week*
    b. *Industry AVG is 4.4 hours/week*
    c. *Laggard is 8.2 hours/week*

By presenting this data to the interviewees, there was a need to understand if the people were aware of the actual spend per process that does not involve time and material. The main purpose of that question was to understand if the interviewees are more focused on ways to optimise a business process than the data that the process uses.

14. ***Achieving BIC status would save you 355hour/year or 8.9 weeks per year how much would you value the time on searching?***

Following on from the question fifteen, it was important to absorb the reaction on presenting these metrics and to engage on more conversation around the process and the quality of the data that is used by the process.

15. ***Regarding Data sources, what kind of actions are in place with regards to mapping out Data infrastructure to identify high priority data sources and what kind of actions are in place to analyse these resources?***
    a. *BIC companies spend 48% on mapping out 36% on analysing*

At this question, the interviewer had to understand how the interviewees prioritise the data sources and if they understand the data lifecycle in general. The main goal was to understand where the focus was. On the process or the data. And if the focus was on the data how do they define any hierarchy within the data sources assuming always that there is a bidirectional integration between more than one system.

16. ***What kind of difficulties do you face during the implementation? And how can you relate to the most common difficulties?***
    a. *Data stored in too many silos usually owned by specific division or department or in ERP or CRM*
    b. *Time*
    c. *Cost of starting the implementation*
    d. *Manual Methods of Handling Data*

Based on the nature that the interview would have taken by now and after discussing the benefits and the importance of the data quality there was a need to identify the difficulties that they had faced so far. The focused is not going to be only on the data quality or the MDM, but on the project lifecycle and how many of these problems are related to data. Keeping in mind that every project consists of the main three elements (people, process, technology) there was a need for further investigation on how the projects, as well as the data, are affected by these three elements.

17. ***How much could you relate to these and how?***

Based on the conversation and the interaction of question sixteen, there was a need to understand how many of the project potential difficulties and challenges were involved around data. That is why on question sixteen, the interviewer had to present some examples of potential issues. The expected outcome of this question was to familiarise the potential data quality issued with the interviewee and to understand how clear is to the process owners as well as the project operators that some of the problems that they may have are related to poor data quality. And the above question would be valid only if there were any problems during the project lifecycle.

18. ***Have you considered buying an MDM product of the Shelf or building one in-house Solution?***
    a. *If the product is off the shelf which one and why and if building one in house solution how did you measure the feasibility and why didn't you buy one off the shelf?*

Moving away from the project lifecycle and focusing more on the MDM as a solution or as a process, the focus should be on the industrialisation of the MDM. The focus was to understand what the users would prefer if they had the option for the implementation of an MDM solution or what was their preference on the solution that they already have in place if they have one.

19. ***Are you aware of the risks of developing an in-house MDM Solution?***
    a. *Usually, it is lack of experience on Understanding MDM, Data governance and access controls*
    b. *The investigation has shown that 49% of the in-house implementations have shown that there was no improvement on the business or the implementation had negative effects -- 10% of the businesses that are implementing MDM prefer non - conventional systems like open source or cloud-based MDM (22% of this 10% is Best in Class)*

Based on the outcome of the question 18, it was important to understand if the interviewees are fully aware of the risks and dangers that may be caused by a false understanding of MDM as a solution if this solution will be built as an in-house project. The main outcome should be the understanding of the preferred option (or not) regarding an in-house development of an MDM solution especially after the potential dangers would have been presented to them.

20. ***Have you predicted the annual data growth? Usually, it is up to 36% but larger companies can show growth up to 75% or even 150% where would you put yourself?***
    a. *40% less than 100000 records*
    b. *29% between 100000 and 1000000 records*
    c. *31% more than 1000000*
    d. *(42% of Best in class companies have more than 1000000 records)*

This question was mainly focused on a combination of answers between questions fifteen to nineteen. There was a need to understand if the interviewees had planned for data growth and what type of data growth, they had planned if any. The expectation was to get the level of understanding from the interviewees with regards to data growth, on static and transactional data and how the data quality would affect that growth. If there was any plan for the data quality growth requirements was the main key point for the desired outcome.

21. ***What about unstructured data formats on MDM? Have you thought about that? and how are you going to facilitate this? Is there a plan to expand unstructured data on your MDM system and how?***
    a. *(pictures, CAD Files, Facebook or Twitter comments)*

This question could not be applied to all industries. However, all the industries that were involved in the research had to deal with big data since all the involved businesses in this research are worth multiple millions. However, unstructured data could exist in different departments of each organisation.

22. ***What kind of Data sources are you planning to use and how many, what kind of integration is it planned? Real-Time synchronous, or Asynchronous?***
    a. *BIC usually use 16 unique internal data sources and 7 externals while all the other companies use 10 internal and less than 4 external*

This is a question that does not apply to all the organisations that are involved on this research based on the specific limitations that have to do with the age of technology as well as the business process requirements. However, there was a need to understand the benefits that interviewees

would from synchronous or asynchronous data integration ensuring the data quality and applying governance on the creation point.

23. In a scale 1 to 10 what mark would you put on the following? (1 low 10 high)

This question was intended to capture the feeling of the interviewees with regards to the part that data play within their specific projects or in their organisation in general. Based on the number that they would insert next to each phrase, with one being the lowest and 10 being the highest, consequently there should be a discussion about the score with the interviewee and try to understand the reason behind their score.

The following table has been given to the participant as part of the questionnaire.

Importance of Master Data Management to Core business operations	
Buy-in from Senior Management	
Resources/ Budget support for MDM	
Adherence to Master Data policies	
Trust in Master Data	
Trust in Data Systems and Policies	
How likely is to identify Records with significant errors	
How likely is to fix these errors in less than an hour	
How likely is for the information to be gained faster	
How likely is a decision to be made faster	
How likely is to integrate new data sources faster	
How likely is for the employee to reduce the time that he spends on data-centric processes	

Table 7 Question 23 Importance of Master Data Management to Core business operations (Gartner 2008)

23.1. Importance of Master Data Management to Core business operations?

There was a need to understand how the interviewees are rating the MDM with regards to their business operations and how the quality of their master data, affect their business process.

23.2. Buy-in from Senior Management?

For each project and organisation, in general, there needs to be a harmonised cooperation between the IT Department which is usually the department that needs to deliver to business requests as well as the business in general who sets up the business processes as well as operational processes. There was a need to understand the views between the three separate groups that were taking part in this interview. These three groups were divided in:

- IT from a development point of view, Business
- Business Process Owners, Senior Management and Decision Makers
- Project Management, Solution Architects and Process Optimisers.

23.3. Resources/ Budget support for MDM?

Even if there was a project of MDM or an element of a data quality project task within the larger scale of a business transformation process and even if there was support for this streamline, there was a need to understand how the business and the decision-makers were supporting the process of data quality insurance and MDM in practice. Bearing in mind that in business everything counts only when monetary values are assigned next to each task or process. That is there was a need to understand what was the level of support by the business on these project tasks is as well as what was the general understanding from the interviewees.

23.4. Adherence to Master Data policies?

Based on this measure, there was a need to understand the level of engagement from both business and general employees that are responsible for executing the business process as well as the IT people. During the conversation, the focus was to understand the processes of engaging in this polices as a project, as a system and as the process. In case of the business transformation project, the interviewer wanted to analyse the “AS-IS” process and the “TO BE” processes and how both sides of the project are adapting on the master data policy.

23.5. Trust in Master Data currently?

Even if it was a BTP or a current data quality/data governance policy that it was already in place, there was a need to understand what the level of trust on the current master data is as well as what would be the trust after a policy exclusively developed for data management including both static and transactional data from both the interviewees but more specifically the senior management and the decision-makers. There was a need to understand that even if there was an investment from the business at some point regarding the data quality management was giving peace of mind to the processes or if there was going to increase the level of trust on the processes after implementation of a project or policy of this nature.

23.6. Trust in Data systems and Policies currently?

Following the above, except for the actual data that is involved in any bidirectional integration defining a business process. It is the system that gives the user the tool to apply and follow this policy as well as to commit any transactions driven by the policy. That is why it was a necessity to understand the challenges that the organisations were facing with regards to systems. This measure was giving two different metrics. If the mark was high, there was a need to understand how the trust has been achieved and if the mark was low, the discussion would focus on the reasons why the level of trust is low.

23.7. How easily do you identify records with significant errors currently?

Based on the mark of this question given by the interviewee, it was necessary to drill down even further. Initially, to understand if it is easy or difficult to measure the quality of the data. How the people were being able to identify a valid record. And what was the business process which defines the identification of a valid or invalid record? At what level they were in the position to identify any discrepancies and what was the impact on both cases. It was essential to understand the impact that it had been caused (If any) based on the time that took the business to realise the data error. What was the impact of a delayed realisation of invalid data?

23.8. How likely is to fix these errors in less than an hour currently?

With this measure question, it was necessary to understand in more detail not only the timeframes that the users need to fix an error but also (in conjunction with the previous metric) to understand the impact of delayed resolution to a potential data quality issue. In industries that are depended on data-heavy processes, a service level agreement with customers is required (internal or external). During the conversation, one of the other metrics that is not recorded was the frequency of these events and how the error resolution had evolved during once a discrepancy would have been identified.

23.9. How likely is it, to get to the required information faster?

This question was focused more on the data management side of the process rather than the service level. The main conversation was around the foundations in place regarding any policies and governance on any integration between systems. How these foundations were assisting in getting the information sooner. The trust in the data and the systems in existing programmes or business transformation programmes should have been identified at the beginning.

23.10. How likely is it for a decision to be made faster currently?

In addition to the above question, there was a need to understand if the information was being used to make any business decisions based on the data. This decision was based on the level of trust that the data governance had been followed and the level of trust on the systems that produce the data. As a result, the information that the systems were producing was complying with the governance and data policies that had been set up. If there were any policies.

23.11. How likely is to integrate new data sources faster currently?

This question was mainly focused on the architecture of the data management process. Based on the metric that the interviewees would be given, there was a debate on how the organisation or the project or the process was bringing together information from multiple systems. If there were specific rules and specific data dictionaries, they should be followed. The people should use

these processes within the existing data ecosystem, or they should take into consideration any reusable processes for future integration between IT environments.

23.12. How likely is for the employee to reduce the time that he spends on data-centric processes currently

The last metric value of the question twenty-three was mainly industry and department related. There was a need to understand how the data is captured by a manual input process, and how the users get the answers to the data related questions that interrogate the systems or the tools that access the data. How the quality of data affects the time that the user spends and how accurate is at the end the process that the user commits.

24. With Regards to Security and Personal Identifiable Information of Confidential Information, what are the actions that have been considered?

a. Anything planned with regards to data encryption (what kind of encryption) of any measures to Data Loss protection to prevent unauthorised access and reduce data loss or exposal

Following the discussions from question twenty-three, there was a need to discuss security and confidentiality of the information that is created, captured, read and used. Depending on the industry, it was necessary to capture the type of information that is clustered as sensitive and how the capture of the information on the data creation process ensures the sensitivity level, as well as the usage of this data once it has been processed. There was also a need to understand the policies around the sensitivity levels and if these policies are mainly followed by the users. If there were any instances that these policies were not be followed, what impact to the business did they cause and how this would be avoided if the polices had been followed or if the polices were there in the first place.

25. With a scale 1 to 10 where do you think that you fit on the following table of Best in Class companies' checklists

For this question, there was a need to divide the metric values into four distinct categories. These categories are the **Process** which deals with how the interviewees see the process of MDM and data quality as an everyday activity. The **Organisation** in which the interviewees evaluate the support from the business with regards to creating the MDM policy as well as the data governance policy. The **Knowledge** which has to do with the information that is available within the organisation and the way that the organisation operates. This division is both data related as well as operational processes as it believed that the department responsible for the data and the overall data management body of the organisation should fully understand how the business operates. Also, which are the standard business procedures, so they can identify any discrepancies or any information generated is not a correct manner. The fourth part is the **Performance** which mainly involved how the organisation measures the performance based on their data usage and data management. These are also the best in class categories for an organisation that wants to implement any data governance or data management policy.

The following table has been distributed to the participants.

Process	Standardized training for master data system	
	End-user needs for data access and use collected	
Organization	Executive sponsor for MDM	
	Cross-functional MDM team	
	Defined Create, Read, Update and Delete (CRUD) role	
Knowledge	Discovery and identification of all business data	
	Classification and definition of all business Data	
Performance	Measurement tools to track and report data quality	
	ROI for MDM defined and Tracked	
	End-Users time to access master data tracked and measured	

Table 8 Question 25 Standardized training for master data system (Gartner 2008)

25.1. Process

- Standardised training for master data system
    - With this metric, there was a need to understand how the training for the process is delivered, if there was any training and how effective this training was. The key element for the training was to investigate if there was a specific

method of delivering this training. In details, if there was a specific training session for specific audiences if the objectives were set before the training again, based on the audience. If there was a different approach with regards the part of the system, the users for each part and if there were different sections assigned to specific users that would make the training more effective.

- End-user needs for data access and use collected
    - Following the above metric, there was a need to understand if before any training was a structured requirement analysis based on the roles of the users. The background knowledge of the users that would take the responsibility of receiving the training but more importantly managing the process in a business as usual basis. Also, it was necessary to understand the roles and responsibilities of each user group that was involved in a master data system or steward a data governance policy on an existing system could fully understand their responsibilities, their boundaries and their influence within the organisation.

25.2. Organisation

- Executive sponsor for MDM
    - With this metric, it was necessary to identify if there was an executive sponsor for the MDM project or data quality process of data governance policy on the existing systems, a new project or the BTP. And of course, the extra important part that the interviewer wanted to understand as if the level of support that the project team or the data management team was receiving from the executive sponsor was the desired one. It was also important to understand the level of alignment between the executive sponsor and the senior management/ board of directors and the level in the hierarchy within the business that the executive sponsor was in.

- Cross-Functional MDM team
    - The main outcome of this question was the roles between the Master Data Management project and how effective and clear the roles were distributed between the team. Also, the experience before each person that now would belong to an MDM team should be evaluated and a “fit for purpose” approach would be assessed based on the level of knowledge of each person that form this team.
- Defined Create, Read, Update and Delete (CRUD) role
    - Training should be given to all the member of the organisation but basically around the data governance and how each department should respect the data that they capture, update, delete or read. It was necessary to assess the level of understanding across all the departments of the organisation, and how each function uses the data, if they are educated enough with regards to the policy but most importantly if these functions within the organisation could understand the value and the importance of this policy. And if these departments were aware of how much their business as usual job is dependent on this policy.

25.3. Knowledge

- Discovery and identification of all business data
    - With this question, there was a need to understand if the interviewees were aware of the term of business data, and operational knowledge. Since the interviewees were covering all the levels of an organisation, it was necessary to understand the holistic view of the business processes. It was important to assess the level of knowledge that these people were holding and how the knowledge they were holding was matching the process that they had to deliver on their project, their business transformation programme, or their business as usual tasks.

- Classification and definition of all business Data
    - The main outcome of this metric is highly connected with the metric just above. Following the previous question, it was important to understand if the users understand the different classifications between each function of the business but also within the same function. Classification of business data refers to a specific set of static data that is used for a specific business process. It was important to assess the level of understanding of this set of data as well as the level of understanding that is involved in the management of this set of data.

25.4. Performance

- Measurement tools to track and report data quality
    - The key metric of this question is the understanding of the importance (by the interviewees) of the quality of the data that they operate within every function of the organisation and every level of their project or programme. The word "understanding", refers to the way that the interviewees measure the quality of the data that they have to work within a daily basis as well as the dataset that they will have to take on, on any new project or programme. That is why it is important to understand how the quality of each data is measured and the results of this measure are used.
- Return on investment for MDM defined and Tracked
    - With this question, it is important to understand from the interviewees (as long as they have an MDM system, process or data governance policy in place) how they measure the return on the investment that the organisation has done on having such a system or process. Or, if the organisation that the interviewee is coming from does not have an MDM in place. What is the return on their investment that they could expect once they have a system like that or a data governance policy in place?

- End-Users time to access master data tracked and measured
    - With this metric, it was important to understand how the interviewees see the benefits (if any) after implementing an MDM solution from an end-user perspective. It was important for the interviewees to estimate the time that it costs the business as usual employees to process the data that they need with the current solution that is in place, the estimation of the time after a solution has been implemented and if the trust on the data that they use reduces the time to complete a process or not.

26. There are four business capabilities and enablers with regards to MDM. Where do you think that you stand and why?

With this question and continue the discussion based on the answers that the interviewees were given, there was a presentation to the audience four different parts that need to take place before any MDM implementation. By presenting these different classifications there was a need to understand what actions had been taken place before the initial phase of the project as well as the actions that had been taken if there was already a policy or process delivered.

This was the table that got distributed to the participants

Baseline	**Organisational Buy-in (Senior Management support, Champion of MDM et)**
	Understanding the existing Data environment
Emerging	Assessing the need for end-users that interacts with MDM
	Measuring Data Quality
Early adoptions	Measuring time to Access Information
	Measuring the ROI of MDM
BIC Differentiations	Standardised Training
	Enforcing CRUD (not define but enforcing)

Table 9 Question 26 Organisation Buy-In (Gartner 2008)

The four different parts are:

26.1. Baseline

- Organisational buy-in (senior management support, Champion of MDM)
    - The first part which defines the foundations of any MDM project or data governance policy is the support of the organisation and especially the senior management or the board which is the part of the business that will assign a budget to it. It was important to understand how the business was supporting the project or programme at the beginning as well as the level of commitment that the users were feeling that the business was prepared to assure.
- Understanding the existing data environment
    - Before a project or a programme of MDM or data governance policy is about to start, an initial analysis of the current environment needs to take place. It was important to understand the actual feeling from the interviewees regarding the data landscape within the organisation was clear from all the departments that would have a significant role during the implementation of this process. It was also important to understand the level of confidence that the interviewees (as well as the business) have on their current data environment as well as how this environment is going to change following the implementation of the new MDM process.

26.2. Emerging

- Assessing the need for end-users that interacts with MDM
    - With this metric, there was a need to define the level of understanding that interviewees had with regards to the usage of MDM from an end-user perspective. It was important to comprehend from the interviewees if the end-users were aware of the MDM as a process, interacts with their everyday tasks. It was also necessary to investigate if the end-users were realising that (even for example when the end-user were performing a business process), the fundamental data they were using to perform this task was based on master

data. How this data was affecting their job and how this data was important for completing this task with success.

- Measuring Data Quality
    - Following the above metric, there was a need to understand how the performance of these daily tasks was affected by the quality of this data. Another part of this metric was to realise if the users were aware of what each selection of the data that they were using was meaning. Even if the level of the data quality was acceptable and within the requirements for performing a task, the end-user knowledge of what this data means could affect any form of measuring the data quality.

26.3. Early adoptions

- Measuring time to Access Information
    - With this metric, the interviewer wanted to understand if any metrics have been taken by the project team or end-users before the implementation of any MDM process or BTP which involved Data Quality tasks. There was a need to understand the effort that has been taken to measure the time and the cost that the end-user needed to complete a task before the implementation and what were the predictions for the same business process tasks or equivalent business process after the implementation.
- Measuring the return on investment of MDM
    - Following the above metric, it was necessary to understand how the recordings that were mentioned on the previous tasks were translated into a return on the investment that the business is planning to make or how the process previous to the implementation was measured with regards to the costs to the business.

26.4. Best in Class (BIC) Differentiations

- Standardised Training
    - With this metric, it was important to understand if the interviewees are aware of the appropriate type of training and how this training plan can be achieved. Additionally, it was necessary to understand from the score that the interviewees would give to this metric the level of thought that has been put into the various aspects of the MDM process, the alignment with the principles of that the data governance policy indicates the understanding of the business process by the end-users that process the data by completing their everyday tasks.
- Enforcing CRUD (not define but enforcing)

• With this question, it was important to understand how the data governance was being applied or strategically defined on how it will be applied when the MDM and data governance policy programme will be live. It was also important to understand by the mark that the interviewees would give if there were specific rules following any new creation, alteration or deletion of data if there was a specific audit trail that was capturing the activity of any alteration on the master data or any prevention of altering this data based on the rules that had been defined by the data governance policy.

27. Which of the following are you thinking of applying or already apply on your MDM toolbox and from 1 to 10 how much that you use/will use?

Since most of the interviewees were on the process of implementing an MDM process and a data governance policy as part of a business transformation programme, the question was more focused on what is in the scope of the project plan that they were working on and a comparison of with their current environment. Another aspect of this question was also the understanding of existing issues that they had on their current environment and how the specific list of metrics would help them realise if this list of actions is included on their deliverables as part of their

schedule or not. This question was divided into three different sections based on automation, data management, and data access.

This is the table that got distributed to the participants:

Automation	**Auto Internal Capture**	
	Auto External Capture	
	Auto Indexing - Sorting	
	Auto Cleansing	
Data Management	Data Enrichment	
	Data Governance	
	Data Deduplication	
	Data Cleansing	
	Data Normalisation	
Data Access	Data Access Tools	
	BI for MDM	
	Internal Collaboration Tools	
	External collaboration Tools	
	Mobile Access	

Table 10 Question 27 Functionality applicable to the MDM toolbox (Gartner 2008)

27.1. Automation

- **Auto Internal Capture**
  - A score was required by the interviewees to describe the level of automation that was taking place during the process of creation, alteration and deletion of data under the MDM underlining process. The internal capture is referring on receiving data from sources that are not managed by the business but the data that these sources generate are used for performing actions of Create, Read, Update and Delete automatically from system to system. An example of this process is when an Electronic Data Interchange (EDI) is taking place between two organisations.
- **Auto External Capture**

- A score was required by the interviewees to describe the level of automation that was taking place from an end-user perspective. The main purpose the question was to understand the level that a system reacts automatically when an end-user tries to Create, Read, and Update or Delete any data that are protected by the MDM platform. In more details, the question wants to capture the way that a system which part of the MDM ecosystem is automatically reacted when an end-user tries to manually update a record which part of the MDM monitoring process is.

- **Auto Indexing – Sorting**
  - One aspect of an MDM solution is the prevention and resolution of duplicate records by identifying the potential candidate record that is clustered as a duplicate. As part of the automation of the MDM process, is the automatic identification of this candidates records and based on a specific set of processes to index them in relationship with the golden records which is the master record.
- **Auto Cleansing**
  - Following the description of the above process, the next metric on the automation of an MDM solution is the auto cleansing of duplicated records or candidates. As part of the automation, it was important to measure the score that the interviewees would give with regards to the auto-resolving process of the potential duplicates. The main outcome was how automated is the process of auto resolving duplicated candidates.

27.2. Data Management

- **Data Enrichment**

- With this metric, it was important to understand how the interviewees understand the data enrichment process as part of a Master Data Management and data governance programme. The data enrichment focuses on the management of the master record and how the process defines the master data as complete. Which attributes describe a complete record and what level of completed attributes are required for classifying a golden record as complete.

- **Data Governance**
  - Following the above metric. The data governance is the part of a Master Data Management implementation that the policy is applied. With this question, the interviewer wanted to understand the level of governance that is enforced on the enrichment of the golden record.
- **Data Deduplication**
  - The data deduplication metric follows the same logic that the question has been described on the automation part of the question twenty-seven but including any manual intervention from the Master Data Management tool users or data stewards or data custodians.
- **Data Cleansing**
  - The data cleansing metric follows the same logic that the question has been described on the automation part of the question twenty-seven including any manual action. In this case, the cleansing activity will include and mainly be happening by the MDM end-users, data stewards or data custodians.
- **Data Normalisation**
  - With this question, it was important to understand how the interviewees were defining the attributes that all together form the golden record. At that stage, there was a need to understand the

process of how a set of attributes is defined and clustered as a golden record.

27.3. Data Access

- **Data Access Tools**
  - With this question, there was a need to understand the way that the users access the Master Data Management platform. How easy it is, or it is going to be when the platform is ready to be used and how good do the interviewees think that the tool will work with the automation.
- **Business Intelligence for MDM**
  - With this question, the interviewer wanted to understand the level of insights that the Master Data Management process is going to give. What kind of metrics with regards to data quality the users will want to see and if these metrics are part of measuring the return on the investment that they are planning on doing with a programme like that.
- **Internal Collaboration Tools**
  - It was necessary to understand if there is any plan for any collaboration tools that will help with data access from the MDM developed by the IT team of the organisation and why these tools would be developed outside the MDM platform.
- **External collaboration Tools**
  - Same as the metric above, it was important to investigate if any tools are going to be bought in addition to the MDM platform and if yes, why this wasn't a part of the initial MDM programme.
- **Mobile Access**
  - With this question, it was important to measure the desire by the interviewees to have access to the MDM platform through a mobile

device and if yes what would be the reason for that and how that could be useful.

28. Which steps do you think that you need to apply to your MDM implementation to improve it? Please select only the ones that you do NOT already have.

This is the closing question of the interview and its main purpose was to finalise the conversation and summarise what had been discussed. The question was focused in five distinct parts of the semi-structured interview and it there was a need to understand from each interviewee what they think that it was missing from their MDM programme and what would they add and why.

This is what it got distributed to the participants:

Secure Senior Management Support	
Implement a Formal MDM initiative	
Implement Standardised training	
Invest in automation	
Enable Remote Access	
Secure Senior Management Support	
Implement a Formal MDM initiative	
Implement Standardised training	
Invest in automation	
Enable Remote Access	

Table 11 Question 28 Senior Management Support (Gartner 2008)

These five parts are the following:

28.1. Secure Senior Management Support

This question was designed to lead to a summary of the view that the interviewee has with regards to the support that the project gets from the senior management.

28.2. Implement a formal Master Data Management initiative

This question was designed to lead to a summary of the view that the interviewee had with regards to the way that the MDM project and data governance policy has been planned from the initial stages to the delivery and how satisfied the interviewees are with this plan.

28.3. Implement Standardized training

This question was designed to lead to a summary of the view that the interviewee has with regards to the training that has been designed for each part of end-users are taking part on the MDM and data governance process. It was important to understand the level of mature thought and planning that had been put into the training process and how important that is based on the interviewees' opinion.

28.4. Invest in automation

With this question, it was necessary to discuss in a summarised way if the interviewees were comfortable with the level of automation that that MDM process was designed. Besides, how confident they were with limited control over managing the golden record if they had any limitations.

28.5. Enable Remote Access

With this question, it was important to discuss in a summarised way if the interviewees were keen to have access to the master data platform through remote access. A subsequent question would be to understand the reason that this is flagged as a need by the interviewee.

3.4. Critical Reflections on Research Methodology and Design.

As described in the two previous sections, this is an overt deductive research project using qualitative methods. Since the main research question is "*How may the invasive circumstances during a BTP be addressed to enhance decision-making with its impact on MDM",* the research design had to ensure that specific variables had to be taken into consideration. The first variable

was the necessity of the environment to be able to support the experimental data collection. As described in section 3.3.1. Research Positioning, a BTP consists of three phases. That is why the experimental data collection should be collected from all these three phases to cover all the potential circumstances that can affect any decisions with an impact on MDM and can also be characterised invasive.

The two companies that got selected as case studies were on a different phase of their BTP and both of them could give an extended insight on how decisions that had taken or they were about to take had affected or would affect their BTP. From the experimental data from the two Audits, an understanding of how different group of people react to similar circumstances. In addition to the data audits, the semi-structured interviews provided more experimental data by having people from different companies, different groups, different business functions and different roles answering the same questions and facing similar challenges but reacting differently and realising the environment that they were operating from a different point of view. This allowed the researcher to act as a social actor to appreciate these differences between people and gather the necessary experimental and qualitative data to develop an outcome and conclude to three artefacts that form a contribution to knowledge.

Methodology	**Reason**	**Contribution**
Philosophy: Interpretivism. •Researcher as a social actor needs to appreciate differences between people, functions, and organisations. •Data Trustworthy and honest. **Structure:** Mono method. Qualitative	With the Interpretivist approach, qualitative research areas such as cross-cultural differences in organisations, issues of ethics, leadership and analysis of factors impacting leadership can be studied at a great level of depth	•Valuable Case Study Evidence for evaluating the decisions made during the BTP on a. Master Data, b. reasons behind decisions made and c. decisions outcome and effect. •Valuable evidence on: -how different roles within the programme reflect and react under specific circumstances. -how each workstream prioritised data-related activities in conjunction with the overall programme.

Deductive	• Starts with existing theory • Research design is structured based on theories in the research field and approaches from recognised sources and institutes.	Following the theoretical best practices this research also identified potential critical areas that may jeopardise the programme if these areas got overlooked.
Primary & Secondary Research	•Case studies & Personal Interviews. •Arguments and perspectives stated in the literature	After merging the theoretical guideline and the practical implication, has been identified that even if the programme directorate follows the best practices, the timing and sequence of executing these tasks may affect the successful implementation and create more invasive circumstances.
Qualitative vs Quantitative Research	•General Framework. This research seeks to explore phenomena rather confirming the hypothesis about phenomena. The instruments that were used were more flexible and iterative style compared to a rigid style. And the methods were semi-structured, focusing on groups and participants observations rather than highly structured methods such as questionnaires, surveys, and structured observations. •Analytical objectives for this research was to describe variation and explain relationships and individual experiences as well as group norms in an open-ended textual question format rather than quantify variations and predict causal relationships describing characteristics of a population on a closed-ended numerical question format. •Flexibility. In this research, some aspects of the study were flexible or had a degree	The structure of the research directs the research to apply the mono method technique such as qualitative methods. Quantitative research method analyses and investigates existing theories after the examination of statistical and numerical data to empower the currently available literature of MDM and data governance (Saunders et al. 2007). Examining the nature of the current research, the research method that is most appropriate to the current study is qualitative research. During this research, qualitative data was collected and processed.

	of flexibility rather than being stable from beginning to end. The interviewees' responses in many instances affected the way that interviews would move on rather than no influence or determination on how to move on. The design was iterative, and the data collection and research questions had to be adjusted to what is learned rather than the research being subject to statistical assumptions and conditions.	
Semi-Structured Interviews	•Characteristics. There was a formal engagement between the interviewee and the interviewer following a specific list of questions in a particular order allowing the interviewer to follow topical trajectories when it felt appropriate. •Circumstances. There was only one chance to interview the interviewees and there were eighteen participants that provided reliable and comparable qualitative data allowing to develop a keen understanding of the topic of interest. The inclusion of open-ended questions provided the opportunity for identifying new ways of seeing and understanding a topic at hand.	The interview questions had been prepared ahead of time of the interviews which allowed the interviewer to be prepared and appear competent. This type of interviews allowed the participants the freedom to express their views in their terms each participant from a different position and point of view, providing reliable and comparable qualitative data.

Table 12 Research Approaches related to Contributions.

Chapter 4: Data Audits & Semi-Structured Interviews

4.1. Introduction

This chapter aims to provide insight into the discovery of the real environment and circumstances of multiple real BTPs. The audits are a representation of what happens in real projects which provided an environment for generating experimentation data.

The first audit occurred in a very challenging environment since the outcome of the BTP was not the desired one. This audit gave a good insight into what went wrong but also how the business handled an undesired outcome.

The second audit occurred in a different environment since the BTP was at its initiation. The organisation acted in a very proactive way. This proactiveness gave a good insight into the initial planning of the BTP and the importance that the MDM and the data strategy overall would play through the programme.

The interviews provided a good insight into how different roles within a BTP react to different challenges and how each role measure the importance of MDM based on their needs and the stage of the project.

The Chapter is divided into two parts. The first part is structured with the two audits in which each audit belongs to a section. The second part is divided into four sections. Each section represents a grouped presentation of the interview answers.

4.2. Audit Reports

Based on the “Data Audit Framework” by Jones, Ross, Ruusalepp (2009) a specific methodology applied to perform the two audits. This methodology is divided into four different sections. These sections are the following:

- **Preliminary Risk Assessment** which includes interviews between the auditor and the senior management of the company that the audit was taking place.
- **Planning Stage** which includes multiple interviews with key employees that currently run the process under audit. The conversations were based on the aims and objectives that came out from the first interview within the preliminary risk assessment stage.
- **Testing Phase** which includes a review in the business processes related to the process under audit. To determine any potential violations of the operational standards or best practices, a set of documents were reviewed and have been combined with the notes that were taken from the interviews.
- **Exit Meeting,** which is the final stage of the audits, represents the wrap-up phase of the audit methodology. In this meeting, the auditor and the company management will review the audit results and discuss any major violations or failures discovered during the testing phase. A set of objectives and actions will be defined based on the findings.

Despite the fact, the same methodology had been applied to both companies, the audit aims and objectives between the two companies were slightly different. The major difference between them was the timing that the audit was taking place.

Company One was performing the audit after a failed business transformation programme and they had already a Master Data Management function within the organisation. The audit that took place in Company One was focused on the performance of this already existed function. While company Two had initiated (then) a business transformation programme and they wanted to ensure that the required controls were in place for a successful delivery.

4.2.1. Company One

The following introduction is a preamble for the recipients of the audit report. The executive summary is an abstract for the C level executive that requested this audit report and it is not an executive summary for the Thesis.

Company One Background

In October 2011, Company One decided to proceed with a complete transformation for both business and IT, by implementing a new Enterprise Resource Planning (ERP) system replacing the two legacy systems, which at that time were both summing up to a 50 year of age. The decision to implement the new ERP system was made just after a new implementation of a new Warehouse Management System (WMS) at a new Distribution Centre (DC).

Due to the successful implementation of the new WMS which was based on Microsoft Technologies, along with the IT strategy (which at that time it was focused in Microsoft Technologies), Company One decided that the new ERP, should be inline (technologically as well as strategically) with the strategy. Thus, the selection was Microsoft Dynamics AX (AX).

Following the market and technological trends, and in addition to the nature of the business, data is the most valuable asset in an organisation and especially for a business - like Company One, the integrity of this asset was the most crucial factor for moving forward as a business and of course for generating revenue. An example of this need was a specific function (Vendor Rebates) in which a significant percentage of the revenue was based on the data integrity of this process.

Understanding the importance of this, the business decided to invest in an MDM solution.

1. There are a large number and variety of systems in the Company One's technological ecosystem interspersed across every sector and department of the group. Products, customers, vendors, and employees are the main datasets that each function uses to perform their tasks and services.
2. All this transformation that was mentioned above, could not happen without populating the new ERP with a proper set of master records. For this reason, a clean, de-duped, "golden" version of the customers, products, vendors and employees' dataset, should be produced from all the incumbent systems and stored in an MDM solution that would ensure this uniqueness and a single version of the truth, for every single record of these datasets.

3. Being a complex and growing organisation, the IT function would continually evolve, creating the need for further integrations introducing new functions, processes, and systems. This evolution brought along new datasets that should be considered a part of the Master Data Management strategy. Having had already a predefined system that measures, manipulates, and merges this data, was fundamental.

1) Introduction

The first audit report took place during a business transformation programme. The organisation was in the process of implementing a new ERP system. Also, Company One was in the process of merging two different businesses. The audit report relates to the desire by Company One to review, evaluate and suggest improvements to the current Master Data Management solution and process that operates within the organisation.

The report articulates the current process and will propose any potential improvements and an understanding of the roadmap and the strategy with regards to the management and the use of master data within the organisation.

2) Executive Summary

The results from the audit showed that there was not an MDM process in Company One.

At that point, there were multiple systems in Company One and there was no articulated and defined architecture for the data distribution.

During the audit process, it was found that there was no MDM process and no centralised solution that acts as a data steward with an applied data governance process. Master Data Management only existed as a concept.

The current state and the future roadmap could not guarantee a single version of the truth and a golden record for the master data. There was a considerable risk with regards to specific processes, due to incorrect product details. Inaccurate product details were also affecting customers and pricing as well as stock and inventory levels.

Customer details were a “considerable risk area” as there was no solution in place managing the golden records, and unfortunately, there were no immediate plans to address this. The same risk applied to vendor golden records. Employees’ records were fragmented and there were no plans to address them. This indicates a risk to the organisation due to the unknown version of the records in all the existing systems.

Regarding the team structure and the investment that had already been made in the MDM department and considering the external consultants that they have driven this practice in the past, it was evident that the root cause of the current state was the lack of structure and communication regarding the departmental requirements, capabilities and processes. The MDM department, at that stage, was employing 19 people, however, none of them had a data management background or training. These 19 people were content creators and data processors. Their role should not demand by them to understand data management, which was a big gap in the strategic roadmap of the organisation. This was evidence of not having an MDM process.

A previous consultant was one of many external consultants that created a practical and feasible plan based on the best practices of MDM. Unfortunately, even then, there was not a hands-on MDM technical architect that could verify the documents technically and implement the vision. The plan was theoretical and made no recommendations of the technologies required to fulfil the ambition.

Currently, the only person within the MDM team that can support the practice from a technical perspective and background is the MDM director. However, a director is not a developer or a technical architect and was not able to technically validate the processes and requirements regarding the practice, especially when there were 19 people to manage with no MDM experience.

To enable MDM practice is recommended that Company One should dedicate a hands-on MDM technical architect and a database developer to work alongside the MDM director to recommend a technical solution. This approach could lead to a design and implementation of an in-house MDM solution or could equally be the introduction of an “off the shelf” solution. Either way, investment into technical analysis and proof of concepts was essential.

A complex organisation like Company One should expect a period of 6 to 12 months to implement an MDM solution.

3) Aims and Objectives of the Audit Report

The aims and objectives of the audit report are only the aims and objectives that had been set for the audit report only.

As it is mentioned above, there was already an MDM policy and process at Company One. The main reason for this audit report was to ensure that the existing process and roadmap was in line with the business strategy and best industry practices. Also, since master data should be one of the highest priorities in the business, a data strategy should be included in a well-defined business strategy. Following that, these were the main aims and objectives that were expected to come out of this audit:

1) What was the current stage of MDM within Company One?
2) What were the processes in place that defined the MDM in Company One?
3) What technology was in place to ensure MDM?
4) Was the Master Data Management practice in line with the Best Practices in Company One?
5) Was there a Roadmap for the MDM solution?
6) If there was a roadmap for the MDM solution, what was the feasibility of achieving this plan?
7) Were the appropriate systems in place to ensure MDM?
8) A review of data flow within the MDM
9) Suggested improvements
10) Suggested Actions

4) Audit Results Company One

During the first three stages of the audit, the following documentation had to be reviewed. The Business Analyst, Director of MDM and MDM Project Manager provided the documentation.

Documentation

Multiple documentation had to be reviewed. The documentation was describing the process on their current format in comparison with the processes in the suggested format. There was documentation for each data stream and each data domain. The detailed documentation can be found in Appendix 2.

Company One Notes from Audit Interviews

1) Then, Company One had two separate functions for their MDM solution. The first path was an MDM solution supporting the business as usual and the second path was an MDM solution supporting the new ERP project as part of the business transformation programme.
2) 19 people were working within the MDM team
    a) 10 of them were content creators
    b) 5 of them were transactional users
    c) 4 of them were ERP Project ETL (Extract Transform and Load), processors
3) Then, the focus regarding the MDM plan on products.
    a) The application that was used for the products MDM solution was Product Information Management System One (PIMS1) (Anonymity) this was based on Oracle technology.
    b) The content creation for products was happening in PIMS1 by the content team.
    c) PIMS1 was the main database for products. Also, there is another application called Adobe InDesign which was used by the graphics department to create the catalogues, which was using the content from the same PIMS1 database.
    d) PIMS1 was supposed to send data to a staging database via BizTalk (Microsoft Integration Tool) and from this database, there were SSIS (Microsoft SQL Server Integration Services) packages that we're loading the data in two versions of e-commerce databases for the websites.

e) From the staging database, there was a tool called DIEF (Integration tool) that had all the ERP rules. DIEF was creating an XML file to get loaded in ERP.

f) ERP was distributing the products that had been imported via BizTalk in WMS.

g) Also, from the staging database, there was a CSV file that was generated. This file then was getting inserted in the legacy ERP1 and legacy ERP2 through their respective importer or manually through an excel spreadsheet.

h) Once the data was in both legacy ERP1 and ERP2, there was a recurring process which was inserting data in the data warehouse databases. From that point, everything was distributed to all the other subsystems.

i) At the time of the audit, a licence free software called DQ1 was used alongside PIMS1 for data profiling.

4) With regards to customer data, at the time of the audit, there were 3 versions of Customer Relationships Management (CRM) which was used for different purposes.

a) There was not an MDM policy or solution supporting this data, but the roadmap was suggesting CRM as the main repository for master data.

b) The data integration into AX was happening by deleting the existing data and inserting the new dataset of customers. Version control did not exist and there were no quality controls except a basic duplicate search.

- The procedure that was searching for duplicates was not including any similarity algorithms or any fuzzy grouping/match. Unless there was not an exact match, any record could be inserted into any system.

5) Regarding vendors, there was a roadmap. This roadmap was based on maintaining the master set of records within the AX.

6) Regarding employees, there was not a roadmap for this dataset. A usual response during the audit interviews was that “it can be left for the end because it is easy and not that important.”

5) Conclusions

Based on the aims and objectives of this audit please find below the auditor’s conclusions:

a) What is the current stage of MDM within Company One?

1) At that moment, there was not an MDM practice within Company One. There was a lot of documentation of what was needed to happen with the correct theory in place but unfortunately, there was not a roadmap based on those documents and best practices.
2) What was happening and what was planned to happen were only integrations between systems.
3) This integration, in many of the cases, was sequential and not centralised, which is the main foundation of the MDM process.
4) There were 19 people in the MDM team which unfortunately none of them was technically educated with the principles of MDM and they did not have the required understanding of data management in general.
5) These people were very good at what they were supposed to do, which was data processing and content creation, but they had not demonstrated an understanding of data or master data.
6) There was confusion within the business of what MDM means. Unfortunately, this confusion was passed on to the MDM department and moving forward the real purpose of MDM was not happening. The business understanding of MDM, in that case, was product creation on a product information management system and catalogues.
7) There was not a defined data governance policy within Company One.
8) There was not a defined data stewardship process within Company One. Based on the interviews and the documentation, the only stewardship that was happening was the code within the BizTalk interfaces. This was not documented and defined and unfortunately it was not in the place that it should be.
9) The only technical person that had delivered an MDM solution in the past (previous roles) was the MDM director, but he did not have the support from the business on delivering per his roadmap. This could be because he was working hard, managing the 19 people on his

team as well as supporting random data requests from the business which was taking most of the department's time. Also, there was not a dedicated development team that could support his vision of delivering his roadmap of MDM within Company One.

10) When the question arose of who had the technical capability to deliver the MDM vision, the answer that the MDM director offered was "himself". However, this could create a huge gap because as a director he should be managing the team delivering his vision, and not being focused on designing, implementing, and delivering a solution which is highly depended on his technical capacity. In this case, the MDM director should focus on managing the very large team of 19 people and be the person who communicates the business requirements, consult business on what needs to be done, agree with stakeholders on what is the next step for both parties to move forward and then deliver this communication and direction to the MDM technical team to deliver the roadmap. (*The above statement is a representation of what has been described in chapter 3.6 of literature review. The MDM director presents himself as the solution to every problem and acts like a "Unicorn" (Woodie 2013).)*

b) What were the processes in place that define the MDM within Company One?

1. At that point, there was not a process of MDM within Company One. Various sources were creating the same data with different references and there was not a central repository for the master data that data governance and data policies could apply. The MDM was happening as an idea on multiple source systems and there was not a referential adjustment within the data ecosystem that could ensure that all the systems were up to date with the latest version of the golden records.
2. Also, data creation was 'one way sequential in parallel' which means that datasets were being created on the source system which was used as a master data repository with some form of deficient governance. Following the data creation, this data was distributed to other systems. These systems then were used as an extra data point for distribution to the next

systems. This process had become a recurring task until the data reach the final stage for consumption.

3. MDM as a process is keeping a single version of the truth of an organisation's master data. Master data in an organisation like Company One, was the master set of products, customers, vendors, and employees.
4. Within Company One data ecosystem, there were multiple sources with different references for each of this master datasets.
5. This method was not in any level MDM as there was not a centralised MDM repository with defined governance and applied stewardship. This repository should maintain a golden version of the data with the appropriate profiles on the data elements and referential adjustment within the historical versions as well as the metadata references from different system identifiers.
6. The described methodology was emerging as a considerable risk for the organisation because many of the key revenue streams were depending on the data quality and the data integrity required for producing revenue.
    a. Customers
        i. There was no master set of customers and all the communications between both parties (Company One/customers) could be in danger of alienated integrity.
        ii. If a customer could change their address, there were multiple systems that this change should have been applied.
        iii. If there was a problem with an order and Company One needed to communicate with the customer, there was a consideration that this communication could not happen in time.
    b. Products
        i. At that moment, the only master set of products was existing in one system and through data integration, was being distributed to all the other related systems.

ii. Changes were allowed and likely to happen somewhere in the middle of the original products which this change could not go backwards and could not be replicated to the next designated systems.

iii. This could increase the risk that wrong products with the wrong pricing could appear on the websites.

iv. Customers could receive the wrong codes and invalid prices.

v. Different products could be stocked in the warehouses.

vi. Products could be "forgotten" with old codes and in wrong locations within the warehouse. There was a risk that they could reach the point of expiration or exceeded their lifetime as well as increased risks of miscounting stock and overstocking products.

vii. There was an increased risk of wrong products being delivered from the warehouse to the dealers/customers.

c. Vendors

i. The most common (and highly important for the revenue stream of the group) problem is the change in the product codes that Company One receives from the vendors.

ii. One of the highest priorities of the group was the vendor rebates. The vendor rebates were based on high-quality reports that were depending on the vendor details, as well as and most importantly, on product sales value, cost value, available stock and sold the stock. All these attributes needed a high standard of product code maintenance. (Therefore, most of the investment on the MDM had been focused on the product management). If the report was not of the highest standards this revenue stream was at risk.

d. Employees

i. At that moment, there was not the appropriate attention and approach focused on the employees' data

ii. It looked like it was overlooked as:

1. Employees are the main users of the systems

2. Employees process information in the system
3. Employees do quality control
4. Employees do the cash collection
5. Employees do credit control
6. Employees do the adjustments on the warehouse
7. Employees are responsible for running the business
8. Employees claim absence

iii. If the right data governance for the employees who performed each action was not in place, then the reward for an employee who performed beyond their set targets, was at risk.

c) What technology was in place to ensure Master Data Management

1. At that moment, there was not any technology within Company One to ensure MDM.
2. There was only the integration between the systems.
3. What the roadmap defined, was that there would be:
    a. PIMS1 (Product Information Management System One) for the product management
    b. CRM for customers
    c. Microsoft Dynamics AX for suppliers
    d. Microsoft Dynamics AX for employees
4. The above indicates that there was not a standard, centralised, bidirectional integration to ensure the MDM process with all the data governance process; elements as well as the defined data policy to ensure the referential integrity.
5. There was not a defined data stewardship process to ensure the data would follow a unified data profiling method.
    a. Data profiling is an analysis of the candidate data sources for a data warehouse to clarify the structure, content, relationships, and derivation rules of the data.

   b. Within Company One, there was a tool called Pandora which was not doing profiling but data quality control and identification of duplicated records (The free version that was used from the relevant department)

6. The dependency in multiple systems to act as master data silos, as well as the ongoing requirement for applying data policies along with data governance in many systems, was highlighting the fact that for Company One to continue following that path, it would have to maintain these policies and processes in multiple systems, multiple times which translates to more time, more personnel, more maintenance, and more risk.
7. This could also mean that the highest level of system dependency arose thus to force more maintenance contracts with the vendors as well as an elevated level of a system upgrade or alteration.
   a. For example, PIMS1 was the MDM solution for products. However, Studio/content creators were using Adobe InDesign for the catalogue creators which was using the same database as PIMS1.
   b. If the studio/content creators were not satisfied with the InDesign and they wanted to proceed to a system alteration (which this was the case), the business would still have to keep PIMS1 system and its data for data creation and a different system for the content generation.
   c. That could mean that the existing licences and infrastructure for PIMS1 would need to remain and different licences and infrastructure for a new system would have to be purchased.
   d. Or to change everything and new interfaces and regulations to be created based on the new system

d) Is Master Data Management practice in line with the Best Practices within the group?

1. Based on all the above, there was not an MDM practice due to the following reasons:

a. There was not a centralised, bidirectional parallel integration with a single point of data governance and data stewardship.
b. There were two different paths for MDM, one for business as usual and one for the ERP project which by accepting this statement, Company One was building two sets of master data.

e) Is there a Roadmap for the MDM solution?

1. There was a roadmap based on multiple presentations.
    a. The presentation documents indicated what needed to be done based on the best practices and theory, however:
        i. What was on the documents would not apply to what was happening and what had been planned to happen.
        ii. The future was indicating multiple systems that would act as MDM platforms.
        iii. These platforms were also the sources of data creation which were invalid based on the theory because MDM should be an independent isolated solution and the only authoritative principle in which the new versions of golden records should be distributed.

f) If there is a roadmap for the MDM solution, what is the feasibility of achieving this plan?

a. Based on the above, the implementation of an MDM solution was impossible, and the roadmap could not support that.
b. What was defined on the documentation, was an integration process between multiple systems? The business should change the title MDM to system integration and data creation strategy.

g) Systems in place to ensure MDM

1. There was not an in-house developed system or an off the shelf system to ensure MDM based on the best practices or even the Master Data Management principals.

h) Review of Data flow within MDM

1. Unfortunately, there was not an MDM solution in place therefore a data flow review within MDM could not occur.
2. The data flow was happening only via system integrations.

i) Suggest Improvements

Improvements could not be suggested at that point since there was not an MDM solution in place.

6) Next Steps

1. Based on all the above, only a high-level suggestion could be given of what was needed to happen to achieve MDM from a technology perspective.
2. From an approach perspective, the suggested actions were as follows:
    i. The business had to understand the importance of MDM due to its nature and to understand what the purpose of the MDM team was.
    ii. The business should support the MDM team by understanding the role of this team or redefine its role.

iii. The business should support the MDM team by adding more technical capable people to support the director and drive the project, rather than expecting from the director to deliver it technically.

1. The MDM director should be there to drive a plan and manage the people as well as to communicate and drive business expectations.
2. The MDM director should not be there to be a developer.
3. The MDM director should demand and make clear his requirements and not always accept without questioning and challenging the business decisions.
4. The MDM director should engage with his plan and roadmap; assuming that this plan was following the best practices for MDM process and communicate with the business the level of engagement that is required from both parties.
5. The MDM director should follow one of the following three options to deliver an MDM solution which was so important to a business as Company One:
    a. Declare and document a full plan that explains what needs to be done regarding MDM, and requests from the business to sign it with names and signatures so that next time that a question is asked of "who requested this process", the answer should not be "The Business".
    b. Choose an off-the-shelf solution that ensures that the MDM process is implemented supporting the business needs and have the vendor to perform a proof of concept (POC) to demonstrate how this would work. Then if the POC would be successful, to engage with this vendor and present the results and ask the stakeholders to engage in a formally signed document. Or decide to develop an in-house solution that will do what an MDM solution is supposed to do following the best practices.
        i. In that case, the MDM director would require two database developers from the development team and train them on the master data services of Microsoft SQL Server.

ii. Supplement his team with two permanent developers who will be responsible for the development and delivery of the MDM Solution.
iii. Hire two contractors; one with MDM architecture experience that could lead the project, and one database developer along with the dedicated project manager. During this process, the director should train two of the employees already on the team to support it.

4.2.2. Company Two

The following introduction is a preamble for the recipients of the audit report. The executive summary is an abstract for the C level executive that requested this audit report and it is not an executive summary for the Thesis.
Details about the System that are involved in the Company's Two environments and their definition please refer to Appendix 3.

Company Two Background

In November 2016, Company Two decided to proceed with a complete transformation for both business and IT, by implementing a new Enterprise Resource Planning (ERP) system replacing the existing ERP system which was also including store till systems (EPOS), warehouse management system (WMS) and product information management (PIMS). The decision for this Business Transformation Programme (BTP) was made after the existing ERP vendor decided to increase the licence fees by 150% annually after an acquisition.
Following the market and technological trends as well as the fact that Company Two operates in both business to business (B2B) and business to consumer (B2C), Company Two had decided to transform their model to a data-driven and data-centric approach with the hope that all decisions would be made based on data.

As part of the BTP, Company Two decided that an MDM and a data governance policy should be applied across the business. Based on that, these are the main key points.

1. There are a considerable number and variety of systems in the Company Two's technological ecosystem interspersed across every channel and function of the business. Products, customers (B2B and B2C), vendors (factories and material suppliers as well as any operational suppliers) and employees (operational staff and in-store staff) were the main datasets that each function uses to perform their tasks and services.
2. All this transformation that was mentioned above, could not happen without populating the new ERP with a proper set of master records. For this reason, a clean, de-duped, "golden" version of the customers, products, vendors, and employees' dataset, should be produced from all the incumbent systems and stored in an MDM solution that would ensure this uniqueness and the single version of the truth, for every single record of these datasets.
3. Being a dynamic and growing organisation, the IT function would continually evolve, creating the need for further integrations introducing new functions, processes, and systems. This evolution created new datasets that should be considered a part of the MDM strategy. Having had already a predefined system that measures, manipulates, and merges this data, was fundamental.

1) Introduction

The second audit report took place during a BTP. At the same time, the business decided to change completely the business model and the way that all the operations had worked so far. The audit report relates to the desire by Company Two to review, evaluate and suggest improvements to the current Data Take On (DTO) process of the ERP project.

The focus during this exercise was to understand the systems that would be involved in the first phase of business transformation which was the development of a Unified Data Model (UDM). A centralised data repository that would combine static and transactional data across

multiple disparate sources into one single version of the various data domains that have been identified during this analysis.

Since the organisation was going through a major business transformation programme which involved a new ERP, a new WMS, and the latest release of an e-commerce platform. The analysis had to take into consideration the new systems that would be the main core of the business operations, and how the data distribution between these systems would ensure data integrity and accurate business Insights.

There was also an element of analysis on the project plan. There was a requirement for analysis on the BTP to understand the sequence of events. The analysis had to be performed on the deployment of each system, the DTO process during the deployment, and the data governance after the deployment.

2) Executive Summary

This audit report combines a system analysis and a soft data audit that took place to define the scope of the DTO process for the new ERP project. The DTO process could be combined with UDM. The UDM would align with the current BTP that Company Two was currently undertaking.

As part of the analysis, a detailed investigation took place involving the systems that are currently used and the systems that are going to be used.

The analysis focused on understanding the data domains/ subject areas that would be involved in the creation of the UDM as part of the DTO. Also, the analysis should define how the UDM could be utilised by the BTP.

Based on the results of the investigation, the UDM needs to run in parallel with the BTP since both depend on established business rules. These business rules should be mapped with an already defined data strategy with pre-defined policies that all the involved project teams follow.

At that moment, there was not a defined data strategy across the BTP, increasing the risk of the BTP itself. Because of the above, many parts of the programme we're running at risk based

on assumptions that the outcome of the programme would be able to accommodate these assumptions.

Except for the fact that a data strategy had not been defined yet, there was also another element of the programme that was increasing the risk. This element was the implementation of a new WMS. The system needed to be connected to ERP to receive orders and send the status of the warehouse back to the ERP. At that moment, there was a considerable risk that the WMS would be delivered earlier than the ERP based on assumptions that cannot be fully and sufficiently tested. Until the ERP system was released, the WMS system would be idle and potentially would require changes close to the go-live date.

The main outcome of the analysis focused on the fact that a data architect or a data officer was needed to be placed as part of the BTP. The data architect would be the person who would take the business rules and built the data strategy based on these rules. That person would also be the one who would define all the data related policies and will ensure that all the project teams follow the same data rules.

Highlighted risks and concerns:

- There was not a data strategy for a heavily data-driven BTP, consequently, there was not any:
    - o Data Cleansing Strategy
    - o Data Integration Strategy
    - o Data Governance Policy
- The project teams were working based on assumptions.
- These assumptions were implemented in silos and not under a specific direction as this direction was not in place up to that point.
- The new WMS was running the risk to become idle for a period due to a potential lack of alignment between the ERP system and WMS releases.

3) Aims and Objectives

The aims and objectives of the audit report are only the aims and objectives that had been set for the audit report. These aims and objects are not part of the aims and objectives of the Thesis.

The main objective of the exercise was to understand and evaluate the data sources that would be involved on the UDM as well as to define the data domains that the business would need to focus on moving forward. However, due to the transformation that Company Two was undertaking, further analysis and soft data audit had to take place to define the final form of the UDM. Also, there was a requirement to ensure that the ongoing BTP would be moving forward on strong foundations leading to successful results that would meet the business operational requirements. Following that, these were the main aims and objectives that were expected to come out of this report divided into two parts:

- UDM Implementation
- Define the data sources that would be involved in the UDM.
    - Define the integration methods between data sources and UDM.
- Define the principal data domains/subject areas that the UDM would be built on.
    - Define the static data that would be a part of the UDM.
    - Define the transactional data that would be used within UDM.
- Define the operational requirements for the integration between data sources and UDM.
- Business transformation project
    - Overall data strategy which involved:
        - Data definitions and data cleansing
        - “New World” data integration strategy
        - DTO Strategy
        - Data governance
- Project deliverables alignment
- Next Steps

For that audit, the UDM implementation and the BTP had to be considered as two separate projects. However, these two projects should be clustered as one with the UDM an important part of BTP for the general business improvement and not only as part of the DTO.

Both UDM and BTP should have one common goal which was business operational excellence and business improvement to effectively accumulate and manage enterprise-wide knowledge based on qualitative data that can be quantified, measured, and improved with a strict focus on return on investment.

From an operational perspective, the reason that BTP was taking place, was because previously (or up to that point) there were technical, organisational, and behavioural issues associated with enterprise knowledge. BTP and UDM together would enable a strategic process that was meant to capture the ways that an organisation integrates its information assets with processes and policies that govern the manipulation of those intellectual assets. The desired goal of knowledge management was the determination and the harnessing of the value of the information resources within the enterprise.

Unified Data Model

During the audit exercise, the focus for UDM was to understand the data sources of the current environment and the “new world”, and how both environments would benefit and use of a UDM. The opportunity to take advantage of the data and information resource could only be enabled if there was an understanding of the structure and knowledge about the collections of information. The critical point was that a formal method was needed for collecting, documenting, and validating business rules. This methodology revolves around ensuring that the information that was present in the system met the expectations of what was in the system. Ensuring data quality was a process of stating information requirements followed by a process of validating that those requirements were being met.

That is why, during this audit, a thorough familiarisation with the key systems and extensive usage took place to understand the business processes.

a) Definition of the data sources that need to be included in the Unified Data Model.

The main systems that would be used as data sources in UDM are the following:

- Main ERP for the new world (ERP2)
- Product Information Management System (PIMS2)
- Retail and Shop system
- Warehouse Management System (WMS1)
- Current ERP which is going to be replaced (ERP1)
- Accounting Software
- Temporary Stock Management system
- CRM
- E-Commerce Platform

The focus was mainly on the transactional systems and the systems that were operating the main processes and information within the organisation. A detailed version of the source systems under investigation can be found in Appendix 3.

b) Define the Principal Data Domains/Subject Areas that the Unified Data Model will be built on.

Building a UDM that includes multiple systems as sources, there is a need to define all the subject areas that will be involved. Subject area or data domain is the grouping of related logical entities that focus on a business area.

For this exercise, the investigation took place on the systems mentioned above and the data domains that have been identified based on the system analysis are as follows:

1. Product Domain

The product domain is an operational-styled hub that manages the definition of products. The product subject area is the main domain that is used across most of the operational processes across the organisation. Each process involves either the product itself or a sub-area. As part of

the product subject area, there is a subcategory of components. The components are part of the production process.

2. Customer Domain

The customer domain is an operational and analytical -styled hub that manages the definition of the customers as well as the classification. The customer subject area is the main domain that is used across most of the order processing operational processes across within the business. In this environment, there are five different types of customers that have been identified.

- The wholesale customer (B2B process)
- The Agent that is similar to a wholesale customer (B2B process)
- The online shop customer (B2C process)
- The retail customer (B2C process)
- The staff that acts like a customer (B2C process)

Each of these customer types has a diverse set of attributes that completes the definition of a customer.

3. Supplier Domain

The supplier domain is an operational and analytical -styled hub that manages the definition of the suppliers. The supplier subject area is used across multiple operational processes mainly focused on Production, services, and receivable goods.

4. Stock Locations Domain

The stock locations domain is an operational -styled hub that manages the definition of where and how the stock is held. The stock location subject area is used across multiple operational processed mainly focused on the inventory measure group of data. Stock locations subject area includes all the dispatch information.

5. Store Domain

The Store domain is an operational -styled hub that manages the definition of the omnichannel approach to retail sales operations. The store subject area is used across the omnichannel operational processes mainly focused on retail sales. As part of the store, the subject area is the store staff subdomain, store stock subdomain and store till registers (EPOS systems).

c) Define Operational Requirements for the integration between Data Sources and Unified Data Model

Regardless of how the data domains/subject areas are determined, the concept of unified data and how this data is identified and managed needs to be consistent across a company's systems and processes. Master, data should be clearly defined and distinguished from or related to other types of data, such as reference data and transactional data. Here are definitions for these types of data:

Master/Static data: Data representing key data entities critical to company operations and analytics because of how it interacts and provides context to transactional data. The sets of master data have been defined within each subject area based on the analysis that has been made at the systems (Russom 2010).

Transactional data: Data associated with or resulting from specific business transactions (Russom 2010).

Reference data: Data typically represented by code set values used to classify or categorise other types of data, such as master data and transactional data (Russom 2010).

Metadata: Descriptive information about data entities and elements such as the definition, type, structure, lineage, usage, changes and so on (Russom 2010).

Each of these types of data will be used together for operational and analytical purposes as part of the UDM deliverables. UDM except the central data repository will be the source of discipline and control over unified data to achieve a consistent, trusted, and shared representation of the Company Two data ecosystem.

The UDM and data governance programmes should work together to focus on managing and controlling the elements, definitions and business processes that influence the creation and change of static data.

Recognising and defining this are perhaps the most challenging and foundational actions within a UDM programme. Especially when this happens simultaneously with a BTP.

UDM and data governance efforts can be and often are initiated with objectives to pull the business entities and definitions together, but this is a much more complicated process that

requires attention from many different areas of the business to ensure a coordinated approach. That is why the UDM, the BTP should perform under a shared and aligned data strategy that includes a shared data governance policy and following the same principles.

To fully evaluate the unified data and the unified data characteristics within a domain, the following objects should be inventoried, gathered, and reviewed for each domain:

Data models: Conceptual, logical and physical models that organize and document business concepts, data entities and data elements and the relationships between them (Russom 2010).

Data dictionary: A listing of the data elements, definitions and other metadata information associated with a data model (Russom 2010).

Functional architecture: Depicts how systems and processes interact with each other within a functional scope (Russom 2010).

Source to target mapping: Describes the data element mapping between a target system and source system (Russom 2010).

Data lifecycle: Depicts the flow of data across application and process areas from data creation to retirement (Russom 2010).

Create, read, update, delete (CRUD) analysis: Indicates where permissions to create, read, update, and delete have been assigned to various groups for certain types of data (Russom 2010). These artefacts are extremely valuable for evaluating the scope, consistency and use of unified data.

These artefacts should also be the basis for defining key metrics that will demonstrate the value and progress for how UDM can drive the alignment and consistency of the unified data across these subject areas. From this type of assessment, current state baselines can be determined, leading to quality improvement objectives that can be implemented and tracked as ongoing quality metrics for each domain. This type of data asset inventory and analysis by domain should also be leveraged to help scope the data governance, data quality management and metadata management practices needed for the UDM and BTP plan and approach for each domain.

Business Transformation Programme

As it is mentioned previously on the document, Company Two is going through a major BTP involving the major business systems. At that moment, ERP1 was the main system that performs the core business processes within the Group. The functionality that currently is covered by ERP1 would be distributed to ERP2, PIMS and WMS1.

Due to the importance of the project, and due to the level of data dependency few elements need to be taken into consideration for a successful delivery. These elements are the following:

- Data Strategy:
    - Data definitions and data cleansing
    - New world data integration strategy
    - DTO Strategy
    - Data governance
- Project Deliverables Alignment

Data Strategy and Projects Alignment

Due to the major impact that the BTP would have on the overall operational processes. And due to the heavy data depended, and data-driven systems that are involved, an Enterprise Data Strategy (EDS) is required to ensure the business continuity.

Data depended on processes that already exist, will have to be mapped to the new processes that will replace them, and both will have to meet desired outcomes of the business strategy. Whilst these desired outcomes to be met, and overall EDS needs to be implemented and run in parallel and being a part of the BTP.

At that moment, this EDS did not exist or at least it had not been identified or presented during the investigation process. However, what had been identified and discussed with members of the team (Head of IT Development and Project Manager) was that different teams

were working in silos without cooperation, as a result, to continue with their tasks based on assumptions that the path that they took (or were taking) was going to be the correct.

A major example of the above is the WMS1 implementation which had been going on for eight months based on the assumptions that the ERP2 data would integrate with no problem with WMS1 but without having defined the ERP2 data yet.

Parts of the EDS should include processes like:

- Data cleansing strategy/policy
- DTO strategy/policy
- Data retention strategy/policy
- Data definition library
- Multiple data dictionaries
- Data integration strategy/policy
- Business intelligence strategy
- And Data Governance Policy affecting all the above and all the processes of BTP

At that moment and based on the audit results during the investigation, none of the above bullet points had been defined putting at risk all the BTP. All the above should be part of the business process review and documentation and should run at the same time.

Since BTP was a data-driven project that executes business processes, the business rules and the data strategy should be aligned and reflect one another.

At that moment, the general feeling that came out of the investigation was that data strategy was something that everybody involved had in their mind but had not planned or had not been aligned with the business rules and any of the project tasks.

4) Conclusions

What did come out of the investigation was the fact that there was not a planned project deliverables alignment with regards to the overall BTP.

The major concern which had been shared with the team was that the two major systems within the BTP were running at a different speed. That is why the WMS1 project was progressing based on assumptions with regards to the data that would receive from ERP2.

The most important concern, however, was that WMS1 would be ready to be used before the delivery of ERP2. This would result in having an expensive WMS waiting for the final release of the ERP2 which the other way should be around.

WMS1 could not run on its own and it could not be linked to ERP1, as ERP1 had its warehouse management functionality embedded. WMS1 needed to have the ERP ready to receive orders, process these orders and report back to ERP2. If the ERP2 was not there, then WMS2 would have to **remain idle** until this process could become available and released within the ERP.

Even if there was an option to link E-Commerce into WMS, (which this was not suggested in any case) E-Commerce would require a data feed from ERP2 and the other way around. Even if this practice was in place, Company Two would have two different systems with completely different data type structures. These two different systems would try to communicate with each other and then be altered to be linked with ERP2 once ERP2 would be ready.

Based on all the above, UDM and BTP should be aligned and both following an already agreed, for both programmes data strategy. This data strategy should match the business rules and should ensure the business continuity having as a target the operational requirements.

5) Next Steps

At that moment, there was an opportunity for these two programmes to create a joint data strategy based on the business rules and to follow this strategy as a leading business objective.

That is why the audit results suggested that a new role should be created that leads the BTP and UDM as part of it from a data perspective. This role should guarantee the success of both projects. The role of data officer should understand the business rules and how these rules are mapped on the IT systems. Also, the data officer should ensure that any data elements involved in these processes were mapped, defined, processed, and evolved; meeting the business

requirements and ensuring that every data journey from creation to reporting is in line with the data governance policy.

Based on the conversations with the team, this role did not exist within the organisation at that time. And nobody within the organisation had the required skills for taking over a task like that.

4.3. Interviews results

This section consists of the answers that have been collected during the semi-structured interviews. As previously mentioned, there are twenty-five semi-structured interview questions which are used as an enabler for a more in-depth conversation, and there is a set of three questions as part of a questionnaire. The questionnaire has also been used as an enabler for more detailed discussion. The twenty-five questions have been grouped into four different subsections within this section. These four subsections are as follows:

- Interviewees' profiles.
- Data quality and master data.
- Master Data Management awareness.
- Data strategy

There is also a fifth part that focuses only on the answers of the questionnaires.

Each subsection is followed by a summary, highlighting the agreements or the disagreements between the different groups and the potential impact that their opinions may or may not have on a Master Data Management implementation. A very detailed presentation of each interview question can be found in **Appendix 4** in which each question is presented by an individual in detail.

Also, each subsection presents the results in a group by group basis.

4.3.1. Interviewees Profiles Questions 1 and 2

Roles: The interviews involved eighteen people with diverse backgrounds and from divergent functions of the business. The roles of the interviewees are divided in 2 different Chief

Information Officers (IT Directors), 3 different Chief Financial Officers (Finance Directors), 1 MDM Director, 3 programme Directors for 3 different BTP, 2 IT Project Managers and 1 Head of Business Analysis. The list of interviewees was completed with 4 Business Analysts.

Project Types and Scales: Due to the nature of the organisations that the interviewees were coming from, the monetary scale of these projects was summing up to tens of millions of pounds. The common focus was on projects that were part of a BTP. The average years of experience of each interviewee were 12.75 years which was distributed on an average of 20 years per CIO and CFO, the MDM Director had 23 years of experience, the Heads of Development were on an average of 17 years. Project directors had an average of 8 years of experience. These 8 years were only the years that they had as project directors excluding any prior project management roles. The same approach had applied to the IT Project Managers which from the average of 6 years of experience any experience before their current role is excluded from the count. The Head of Business Analysis had 4 years of experience as head of the function and 13 years as a Business Analyst. Also, the 4 different Business Analysts that were interviewed had an average of 4 years' experience per analyst.

Summarised profiles of interviewees:

Role	Number of Interviewees	Average Years of Experience per Role	Function
CIO	2	20	IT
CFO	3	20	Business
MDM Directors	1	23	IT/Business
Head of Development	2	17	IT
Business Transformation Project Director	3	8	Business
IT Project Manager	2	6	IT
Head of Business Analyst	1	4	IT/Business
Business Analysts	4	4	IT/Business

Table 13 summarised profiles of interviewees

4.3.2. Data Quality and Master Data, Question 3 to 5

There was an expectation due to the nature of the projects that the involved participants would answer that data quality is the second most important part of a BTP. The most important part would be the business process definition and delivery. Out of the 18 participants, 6 people were purely IT, 6 people were purely part of the business function and 6 people were cross functioning between IT and Business. The anticipated answer would be that for every BTP the data quality, should be excellent since the potential alteration of a business process, ensuring the business continuity. However, this was not the case. Except for the people in IT, the function that agreed that the data quality should be extremely important due to all the integrations and all the technical processes were depended on this data. The opinions on the other functions were varied.

4.3.2.1. Interview Answers Questions 3-5

The CIOs (IT)

Both IT Directors answered that the quality of the data is extremely important for them and their function. One of them mentioned that it is extremely hard to ensure that the quality of the data is kept to the highest level.

The CFOs (Business)

All three finance directors focused on the importance of data quality on reports. One of them stated that “It is very important to know that the report that I have in front of me is accurate and that I can trust it” (Appendix 4). The common part of their answer was the fact that they need to know the financial state of the business at any time, so they can make decisions. Each process of the business needs to be constantly evaluated so the cost will always remain low and the profit will be optimum.

MDM Director (IT/Business)

As expected, the MDM director embraced the importance of data quality. As the role indicates, ensuring that the quality of the data is one of the most important daily tasks for him and his team.

Heads of Development (IT)

Both Heads of Development said that the data quality is very important as every application that is delivered is based on data exchange and data integration. If the data is of poor quality, the process that the applications are executing will fail and "IT will take the blame" (Appendix 4 question 3).

Business Transformation Programme Director (Business)

The BTP Directors' answers were not as expected. Two of them recognised that the quality of the data is important but their "job is to make sure that the project will be delivered" (Appendix 4 question 3). The other one though stated, "We are not so concerned with taking data over from ERP1 (This is the Company Two BTP Director from the second audit) if it is inconsistent to start with" (Appendix 4 question 3). Also, two of the programme directors had data transition and data quality on their programme plan and the third one had the data transition and data quality as last part (after 14 months on the project) without any plans on data strategy in general.

IT Project Managers (IT)

Both Project managers understood the importance of data quality and both said that to have "a successful project delivery you must ensure that your data is correct" (Appendix 4 question 3). When the conversation led to the question of how they ensure that the data quality requirements are met, one of them answered: "This is a business activity" and that "the business needs to decide".

Head of Business Analysis (IT/Business)

The Head of Business analysis stated that he understands the data had to be correct, but their role was to ensure that they write down the "AS-IS" and "TO BE" process in time as requested by the Project Manager and not to do data analysis. During the conversation, a question had been asked on if their "AS-IS" and "TO BE" process definition included data definition and rules for data integration, both answered "No".

Business Analysts (IT/Business)

Out of the four business analysts, only one stated that the data quality analysis and the data definition before business process description is necessary because this is how he would be able to define the process by "visualising the data journey". The other three suggested that the data

(not mentioning the word quality) is important, but they should spend most of their time documenting what the end-user was going to do and they should not do any data analysis as this had not been asked by their manager.

Master Data

With regards to the classification master data, both IT directors stated that they understand what master data was as well as all the financial directors. The MDM Director had not been asked if was aware of the classification master data. All Heads of IT development, Project Managers, Business Analysts as well as the Head of Business Analysis and the BTP Directors stated that they were fully aware of the term master data. (One of BTP asked, "You mean the static Data?" (Appendix 4 question 4))

Definition of Master Data

During the fifth question, the participants suggested some interesting definitions of what master data is. Almost every participant had a different definition, however, every definition could be related to the following definition which is an attempt to bring together what all the participants described. Based on the combined answers, Master Data Management is the process that runs in parallel with any business process. Any people that run this business process and any technology that enables this business process, and used by the people to capture, maintain and define the accuracy of a data record that belongs to a specific or multiple data domains, through the data lifecycle following a defined data governance policy.

Both **IT directors** defined Master Data Management as "The tool that manages and maintains the Golden Record" (Appendix 4 Question 5).

The **Financial Directors'** definition was: "A clean version of customers". There was an additional question on what they meant by "Clean" the answers were mainly a de-duplicated version of the customer which enables accurate reporting.

The **MDM director** defined Master Data Management as "Managing a single version of the Truth" (Appendix 4 Question 5). His focus, however, was on the content of the products. That is why his definition was not focused on the customers and any other data domains.

The **Heads of Development** defined master data as everything static and that needs to be managed through a slowly changing dimension. By slowly changing dimensions the Heads

defined the sets of static data that an update or deletion needs to be captured for accuracy on reports.

All the **Business Transformation Programme Directors** defined master data as a "Clean set of Customers, Suppliers and Products". The Program directors, focusing on BTP were referring more to a DTO (Data Transition from the old systems and processes to the new) strategy.

All the **IT project managers** defined master data as all the static data that are going to be used on their projects.

The **Head of Business Analysis** defined master data as the set of Customers, Products and Suppliers that are going under a profiling exercise to be cleaned.

The **Business Analysts** following from the Head of Business Analysis, defined "Master Data [as] the set of Customers, Products and Suppliers that are used for the process definition. This Data needs to be cleaned and de-duped." (Appendix 4 Question 5)

4.3.3. Master Data Management Awareness Questions 6 to 8

4.3.3.1. Interview Answers Questions 6 - 8

Importance of MDM

Both **CIO**s declared that the MDM is very important for their project since they cannot deliver projects where the static data are not maintained. Also, the only way to maintain the data is through a data governance policy that must be understood by the business. If the business will not follow a data governance policy that forces their employees to comply, the MDM will never deliver the expected results. Therefore, every project delivery that is undertaken by IT will always have problems and would lead to more expensive maintenance and manual fixes to the problems.

The **CFO**s had a different view on the importance of MDM. They both focused on the accuracy of the reports and the cost of maintaining this accuracy. They identify as the most

important part of the data management (in general terms and not only master) to keep their insights accurate and their cost low.

According to the **MDM Director**, MDM is very important. This process is "Business as Usual" for him and his team. "The content of the important attributes to define a product is important to be accurate as well as the single version of the truth for the products. The business process needs to ensure that that list of products that I promote to the business is clean and with no duplicate records. Once I sent the data to the business the list is used for many other operations. Except for the E-Commerce website that needs to be "spot on", I need to make sure that the products are clearly defined in the warehouses, so the business will know exactly what levels of stock they have at each time."

Both **Heads of Development** understood the importance of MDM and embraced the importance of the process for their projects. At that moment, they were facing problems with development that is due to the substandard quality of data. One of them stated: "Poor Data in; Poor Data Out". In most cases, the delivery of the project is at risk due to the lack of quality data. The quality of the data indicates the successful delivery of a project or not.

All the **BTP Directors** understand the importance of master data on their deliverables, but they do not include a plan for MDM as a process or maintaining the data quality. They state that is more important to deliver the programme and define a data strategy after the delivery. Also, they explain that it is difficult and time-consuming to go through the business users and ask them to take time out of their daily tasks to define the "AS-IS" and "TO BE" process. They would not initiate data governance at the same time because that would be challenging their deliverables. One out of the three participants stated, "If the data is of substandard quality, we are not taking it into the "new world". We are focusing on the new process and the old data can come later!"

The **Project Managers** declared that the MDM is important, but if it does not exist as a deliverable objective, they cannot do anything about it. The business needs to decide at the beginning to plan and budget an MDM tool or process and if this is approved then they would initiate a plan that includes an MDM process.

The **Head of Business Analysis** identified as MDM being an important part of the process, however, if this does not exist as a priority by the project managers, they cannot spend time on business analysis and any analysis in general that is not part of the general roadmap and plan.
The **Business Analysts** stated that the data needs to be "clean and tidy" so the processes that they design could be effective. However, they are more focused on defining the process as it is now and how is going to be, rather than planning on ensuring the quality of the data that is required. There was an assumption that the quality of the data and the master data, in general, is included.

Best in Class, Industry Average and Laggard MDM

The **IT Directors** were not aware that an MDM can be characterised in three distinct categories, but they assumed that there would be something out there to define the best in a class method. They suggested that it would be ideal to reach the best in class status. However, they both suggested that at that moment, there was not an established process to represent MDM. However, they were making a slow process to improve the management of their master data.

The **CFO**s were not aware either that an MDM process can be classified, and they never thought about "putting a label on the way that we are managing the Data" (Not specifically for master data but for data in general).

The **MDM Director** suggested that he could classify the MDM process based on the data completeness and the level of "cleanliness" of the data. However, he did not mention anything about the process in general. The focus was the content of the data and how accurate this data is.

One of the **Heads of Development** suggested that the best in class classification should capture the data on creation or alteration and validate on the master data silo. Then based on the data validation, the transaction could be committed. However, the level of systems that the organisation had in place would never allow this kind of process to be implemented as an extra step. The other Head of Development stated that they were trying to capture the data on the

creation and that they have a comparison process in place. This process is initiated before the record is inserted into the destination. However, this process is only applied to limited systems. Both seem to be aware of the classification of the process and they would put their current system to a lower level than Laggard.

None of the **Business Transformation Program Directors** was aware of the classification that can be applied to the MDM Process.

One of the **IT Project Managers** as part of a project in a previous role that had to do with the creation of a Centre of Excellence for Data Management, and he had come across the classification. However, his focus was to manage the budget most effectively. The Centre of Excellent for Data Management was at the initial stage of establishment and he did not have to get involved in the deep details. The other IT project manager was not aware.

The **Head of Business Analysis** was not aware of any classification with regards to MDM, however, he thought that there would be specific Key Performance Indicators (KPIs) that separate the processes from optimum to underperforming.

None of the **Business Analysts** was aware of any classification with regards to master data management

Awareness of MDM KPIs

Both **CIO**s declared that there was not a formal MDM system in place in their organisation and that there was not any budget assigned to implement one at that stage. Both scored the question with 1. Both said that they are tasking the IT to follow specific rules on the creation of the data on the systems that they use (Service Level Agreement (SLA) systems and incident management) and the systems that they deliver to the business. However, except for very few occasions, everything is manual. Within the organisation, the rules are easily overwritten. They scored the question with an average 2. (The first one measured with 1 and the second with 3). For the third part of the question, the one CIO stated that there was a collaboration with the MDM team - Content Creation using PIMS1- but there was not any collaboration with the other parts of the business. The other suggested that there is very limited collaboration between the business and

the IT and in some instances, there were different points of view on what needed to happen. The average score for the third part was
2.5 as the one scored the action as 1 and the other as 4.

All the **CFO**s said that there is not an MDM system in place and that they understand that this should be something vital that they can apply a form of governance. Both suggested that they should be the ones that they should support this process and that they should ensure that a project like that should be on the near plans. Following that, the first part had an Average score of 1 as all the CFOs marked the question with 1. The same score was given for the second question however, all of them looked confident that the future systems will automatically force specific rules on the data capture. For the second question, they scored the current state with 1. However, the future state is got scored with an average of 6 having one mark it with 8, and the other with 4 and the last one with 6. The latest score applies for the third question as all the CFOs appeared to be optimistic that the new world would enable better communication and collaboration between the business and IT. However, for the current state, the score was 4.66 as two of them scored it with 4 and one of them with 6.

The **MDM Director** stated that there was a formal function of MDM but there was not a system in place. This was his plan and his active project. At that time, he was going through a due-diligence process with three vendors and as soon as he had any implementation and licences costs, he would take it for discussion to the decision board. His final mark on the Master Data Management system in place was 1 due to the current state. The second part which had to do with the automated capture of the data on the creation scored with 1. Even though there was a dedicated function within the organisation which deals with the master data, the business was limiting the liberties of this team in managing product content only. The creation of customers or suppliers was out of control. The business was operating with 4 different ERP systems and 4 different CRM systems and it was unmanageable. Even though the business has budgeted an MDM function, there was not any commitment in applying any governance on the creation of the records as they were “too busy” to deal with that. That lead to the next part of the question which he marked it with 1. The business had a separate group of people that dealt with data due to the ERP consolidation project. This team were independent of IT. They believed that this was

the correct process because this is what they used to do with the old systems and that IT should deliver what they were asking for. They thought that they were better fitted to capture the requirements, even though the IT was suggesting otherwise.

All the **Heads of Development** scored the first part of the question with 1. They stated that there was not an MDM solution in place and they suggested they cannot put any KPI, as even the processes on the new ERP implementation was not included on the overall solution design. The discussion focused on the new world and how the business did not consult the IT in principal to adjust the issue of master data. As a result, even though they were aware that data from multiple domains were created in multiple different systems, the new world would maintain the same issue as customers were still going to be created into CMR and ERP as different entities and product could still be created into ERP and PIM system. The score from all the Heads of IT Development was 1 on the second part of the question. Regarding the third part of the question, the Heads of IT Development gave a score of 1 as they were suggesting that there was a disconnection between the IT the project team and the development in general regarding MDM at that point.

All the **Business Transformation Directors** scored this question with 1. The reason was that even if they were going through a process transformation, they did not have a separate system to act as an MDM solution. However, they were confident that the new system would have all the required functionality to ensure “clean data”. They also suggested that it was the standard functionality of the system. Also, within the new business transformation ecosystem, multiple disparate systems allowed the creation of multiple records under the same domain. Based on the above, there was a question on how they could ensure the creation or the alteration of the same information in different systems is aligned. One of them suggested, “that the new systems will take care of that process”. Regarding the second part of the question, none of the three programme directors could understand the reason behind the existence of this practice. When it was explained to them, they were insisting that the “new world will not allow duplicates and will clean the data that will come from the other applications”. There was a question regarding the mark they would put on the status and the future status of this process. One of them scored 7, the other one scored 4, and the last scored 6. At that point, there was a question regarding the

high score in a process that currently was not there. They answered that "in the 'new world' the system will have a data cleansing process which will take care of the duplicates". On the third part of the question all the participants answered that for each business function that would have to manage static data, there would be a person responsible to manage this data. This maintenance would be happening through "the system" and if there was a specific request IT would be involved.

The **IT Project Managers** marked the first part with 1 as on their projects there was not any formal MDM solution in place. One of them had asked about an MDM solution for managing the static information, however, there were not any immediate plans at that moment. The project plan indicated that this should go under the DTO process. There were specific people responsible to validate the content of the static information through the system. The other IT project manager mentioned that on all the projects that he had worked, the creation of a data dictionary for the static information was the most time-consuming part and that there was a detachment between IT and business. IT could provide the attributes, but the business could not validate the content. Also, the data mapping task was happening at the end of the project instead of the beginning. That was why the other IT project manager suggested that a formal MDM solution or "something that creates an index for the content between the business and IT should be in place" (Appendix 4 Question 8). However, both marked the formal MDM solution with 3 as it was not in place. When they realised the importance of it, they started managing it manually. For the second part, both stated that there was not any plan for automated data capture at the creation, which is why they marked the second part with 1. For the third part of the question, both project managers suggested that there was not a cross-functional team in place, but the plan was that one dedicated person from the business function would be responsible for managing the static data and when there was a problem would inform IT.

The **Head of Business Analysis** suggested that there was not any formal MDM solution in place and "at the moment the main focus is the definition of the processes" (Appendix 4 Question 8). However, he suggested that through the process there was a part that had to do with the data quality. When the Head of Business Analysis was asked to describe in more detail the part that data quality was involved, he mentioned that the DTO of the customer from the existing systems

into a new system had to follow a cleansing process before insertion as well as products and suppliers. A question was raised about, what would happen after the initial insertion, the Head of Business Analysis suggested that the new systems should be able to take care of any duplicates. Since there was not any Master Data Management solution in place, he marked the first part with 1. On the second part of the question, the Head of Business Analysis described the current process of automatic data capture as very important for the business but "at the moment is incomplete" there was not a validation of the data that was automatically being captured and that in the new world, this process would still be incomplete unless the business decides to follow a specific governance or policy. However, the new systems, would be in position to manage the different versions of the updated data. Regarding data creation, the systems would be able to capture the data on the creation but would not be able to validate what has been inserted. The constraints within the systems would be very limited and would be looking for exact duplicate records. In this case the system would not allow the data processing. The Head of Business Analysis marked that part of the question with 4. On the third part of the question, the Head of Business analysis stated that there would not be a cross-functional team between the IT and the business to manage the data. The business would assign specific roles to specific users to act as data stewards in a "potentially" system that will be provided by IT; however, this system had not been planned yet and these users would be managing the data through spreadsheets. The mark, in this case, was 4.

All four of the **Business Analysts** marked the first part of the question with 1 as there was not any formal MDM system in place and there was not any plan for any separate system. The business focus at that time was the delivery of the BTP. Their answer to the second part of the question was varied between them as two of them marked the automated data capturing with 6 and one of them with 4 and the last of them with 3. The two analysts that marked with 6 were suggesting that the new process would capture all the product, customer, and supplier data on the creation through the system. If there was a duplicate record within this system, the creation of a new record would not be allowed. However, when they were asked what would happen with the other systems that the creation of the same item was allowed, they stated that the process would not have a constraint on the other systems but only on the main ERP. The same

conversation happened with the other two analysts and the same answer was given. The capturing and the validation of the records was happening only on the system that the user was trying to insert a record and there was no validation across the systems. However, all marked with 5 the third part of the question regarding a cross-functional team between IT and Business. They suggested that there was not a cross-functional team to manage data but there was effective communication between IT and the team that manages the data from a content/business perspective. Whenever there was a problem IT would help them to resolve. Another synergy between these two teams that came up from one of the analysts indicated that once the team that was responsible for the content which was maintained in an excel spreadsheet, they would pass this spreadsheet to IT for importing it into a different system.

4.3.4. Data Strategy Questions 9 to 22

4.3.4.1. Interview Answers Questions 9 - 22

Data Strategy

In this question, both **CIO**s were aligned with regards to their data strategy. Their focus was on three things. data quality, ease of access and data archiving. Their data strategy had been structured against the fact that the information should be accurate, easily accessible for meaningful insights and easily compared archived data. Their data strategy should follow their business objectives to allow the organisation to get accurate insights, so the business could make informed decisions.

All three **CFO**s had one common goal. This goal was the growth of the organisation. The goal was the same across all the CFOs and their need was “Accuracy”. All three of them wanted their strategy to focus on the accuracy so when their teams providing them with the necessary reports, the results should be correct. The most usual challenges were that “One report says £x and the same report that comes from a different department says £y”. The main problem that they wanted to resolve with the data strategy, was the elimination of different silos within the

organisation that manipulate data. There should be an enterprise-wide data strategy that would enhance the links between divisions and create consolidated business insights.

The **MDM Director** had a different approach with regards to data strategy in comparison to the CIOs and the CFOs. His focus was around "People, Process, and Technology". He described his strategy as a holistic view of the data across the organisation, but the most important part of the strategy was the data governance. He mentioned that to create a data strategy, the organisation had to follow specific rules and most importantly to comply with these rules. If the end-users followed the rules within each department, the data strategy would be a valid process. By the time that the policy was over-ruled, the strategy did not affect and it would not be valid.

Both **Heads of Development** suggested that their main point regarding data strategy was all the applications and integrations to have embedded rules that were aligned with the process. Their main concern was that most of the times, the business assumes that these rules were known to the IT and when a problem could arise, the IT had to re-focus on fixing the problem without ensuring that the process had been explained correctly. The data strategy based on their opinion should be described in detail all the rules required to avoid any misconception in any application or integration development.

The **Business Transformation Directors** had a different view with regards to the data strategy. They all thought that the business should provide the transformation programme with a data strategy and that their responsibility should be to apply it within their programme. With regards to how important was the existence of a data strategy or specific rules and policies with data on their objectives, they all suggested that they understand the importance of the data rules. However, these rules should come from business. With regards to data governance within their programme, one of them suggested that it should be managed from the software vendor and it should be reviewed by the business in a later stage. There was a question with regards to new software releases, and how this was being managed without specific rules, the answer was "we are more focused on the way that the new system should work in the way that the users are supposed to use it". Then there was a question regarding any user errors when someone does not follow the correct process, the answer was "the system will be delivered in a way that will not allow users to use it differently than the way that it is supposed to work."

Both **IT Project Managers** suggested that they did not have any documentation with a specific data strategy. The only rules that they had to follow were the process within each IT deliverable that they had to follow. Their focus on each project was to understand the input, the transformation (if there was any and the output). This process definition was something that they should take from the business analysts and as long as the business analysts had defined the "AS-IS" process and the "TO BE" process it should be included on the project deliverable.

The **Head of Business Analysis** was in line with what the IT Project Managers were supporting. The head suggested that there was not a specific data strategy and that the only aspect that they took into consideration with regards to the data management within each deliverable was the "AS-IS" and the "TO BE" process and how the data within the process were defined on the input and the output.

The **Business Analysts** followed the same approach with the Head of Business Analysis and the IT project managers were the translation of the "AS-IS" and the "TO BE" process. What they suggested was the fact that when a business user was explaining them the "AS-IS" process and how this process should be altered to the "TO BE" process, they did not follow specific rules. The "TO BE" process did not follow a guideline based on a data strategy. It was also mentioned that in most cases, the business users would disagree with the fact that the new process was there to optimise the current operation, as a result, of the "TO BE" process being a more complicated version of the "AS-IS" without taking into consideration any new data validations or rules that the users should follow.

Senior Management Support

For this question, the **CIO**s' answer was interesting since the role of a CIO is a senior role. However, both of the CIOs said that they support the direction of an MDM solution, however, they find it "difficult to technically convince the senior management of the other business functions." For example, one of the two CIOs stated: "All the board understands that there is a risk with the data within the organisation but the budget approval process for a project like that

is very difficult." (Appendix 4 Question 9) When it comes to budget CAPEX (capital expenditure) for a project that involves various functions of the business, there is always a debate on which business function should take the cost. This is when all the projects start reducing the feasibility. "In this organisation, the CAPEX approval committee meets once a month to approve projects that the estimated cost is above £x thousand. Usually, the operational challenges of the organisation go as P1 (priority one) and at the end of the appointment, a PowerPoint presentation needs to take place for each project. If this presentation is too technical, they will request a revision and a presentation on the next month's committee meeting. The next month, if the revised version of the presentation is simple and lack of details, they will request again more details for the next month's meeting. At that time, there have been 3 months trying to convince the board that they need the project based on a PowerPoint presentation as well as a usually 3 months' work before the first presentation for analysing the environment the requirements as well as the costs. At that point, we have spent 6 months trying to convince the board that they need to step up and support a project like that bearing in mind that all the challenges that are discussed on each meeting have to do with data quality." (Appendix 4 Question 9).

All the **CFO**s have agreed that they would offer all the support that it is required on projects that will help the organisation to make better decisions and improve the business insights as well as the operational activities. One CFO mentioned that "I always support improvement on the process if this improvement is going to reduce the time, effort and cost of the current process without any quality reductions on the outcome of the process." (Appendix 4 Question 9).

The **MDM Director** supported in a very direct way that the senior management does not support in practice the MDM improvement. He suggested that even though he has a very large team that they are supposed to do MDM, what they do is content management via Excel. They do not use any sophisticated tools to help them understand and identify any incorrect records and they are cannot force a data governance policy across the organisation. The MDM Director suggested that every time he went to a commercial function of the organisation to suggest improvements on the way that new customers were created, the commercial function replied

with that they understand that the process in place is not appropriate, but they do not have time to change it. “They are happy to talk about it but not change it” (Appendix 4 Question 9).

Both **Heads of Development** suggested that senior management does not support the MDM in practice, but they support it in theory. The senior management realises that most of the challenges at an operational level come from substandard quality data. However, when the discussion comes around the implementation of a solution that would reduce or resolve these issues, there is a significant difference between what they discussed and what they are willing to do. Even though, on any piece of application that they deliver as the IT Development Department, they are trying to maintain high-quality data. Based on the circumstances, they cannot apply constraints on existing systems that limit the users’ ability to enter inaccurate information on these systems. They cannot prevent the “Bad Data In – Bad Data Out” habit that existing systems allow at that moment.

On the other hand, the **Business Transformation Program Directors** suggested that the senior management supports MDM improvement. That is why they invest in new systems and redesign/transform the way that the business operates. However, when referring to previous conversations when they mentioned that MDM is not part of their current project, (as the data quality and the management of the static information is handled by the new Systems), they mentioned that an MDM system is not in scope at that moment, but the maintenance of static information is something important and that the senior management supports this process.

Both **IT project managers** described the level of support from senior management as inconsistent. Within the multiple functions of an organisation, you have some directors that support the MDM as a process and as a “Must Have” but there are others that do not understand the value. Usually, they end up with mixed messages in which they support the process, but they cannot follow it.

The **Head of Business Analysis** suggested that the support from the business can be measured on the fact that when a project regarding the improvement of a process starts, it requires a definition of the business process. The implementation of this process always depended on correct and accurate data. The fact that there is no specific data governance in

place as well as not constraints to guide the user into entering the correct information is because the senior management is not prepared to follow their own rules.

The **Business Analysts** suggested that there is support from the senior management but there are always time constraints which do not allow the suggested plan to be completed and followed as it has been scheduled at the beginning.

Data Consistency

The solution that both **CIO**s suggested as the current one for managing and maintaining data consistency is a centralised data warehouse that integrates with all the systems. Within this data warehouse, the dimensions that define the schema of the data warehouse usually holds the static information and within the dimension tables, there is a process that can identify duplicate records. However, this means that transactions have already been made with inaccurate data and then through manual intervention (by the developers and hard-coded rules by the developers) these inaccuracies are usually are eliminated for any reporting. Once these inaccuracies are identified, the IT Department will inform the business function to manually adjust the records that contains inaccurate or duplicate data.

All 3 **CFO**s said that the main problem that they have is what has been paid in and what needs to be paid out. This is something that at that stage happens manually by the dedicated departments' accounts payable and payment collectors. However, the reporting is happening by Business Intelligence tools that are attached to the data warehouse database that is being maintained by the IT Department.

The **MDM Director** stated that the only tool that his team uses, is an Excel spreadsheet per data domain. Once this spreadsheet has an updated version, it is sent to IT where they upload it to the relevant systems and the data warehouse. There is nothing sophisticated as it stands at the moment with the current process.

The **Heads of IT Development** described the tools that are used to maintain a data consistency except for the integrations that are happening between the systems, they receive an Excel spreadsheet with static information that will need to be imported after has been evaluated by the business. There is a process that executes specific tasks which process all the data to each

system and the main data warehouse. If there is a duplicate, then the development team will have to understand the reason this duplicate exists. Most of the times different functions of the business direct the IT team that there are duplicates for a reason. This reason according to IT is not valid but they cannot go back to the Business telling them that there should not be a duplicate. Except for the automated integrations between the systems and the Data Warehouse, everything else is the manual investigation. The duplicates are identified only if there is a perfect match on a combination of attributes.

The **Business Transformation Project Directors** suggested that the new environments are going to support data consistency maintenance as a process within the systems. However, all the existing data that needs to be imported into the new systems will have to come from IT in an Excel format. This Excel will have to be manually validated by the process owners before inserted into the new systems.

At that moment, there was nothing sophisticated in maintaining data consistency across the systems. Both **IT project managers** agreed that everything is manual and is maintained in an excel spreadsheet. Once this Excel spreadsheet has the last version is distributed to IT which will import it into the relevant systems.

The **Head of Business Analysis** suggested that it is very difficult to maintain, manage and ensure data consistency across the systems. Everything is happening manually on an Excel spreadsheet that is going to be used as the master file of static data to be inserted on the relevant systems. However, once this set of files will be inserted into the new system, the designed future process will maintain the accuracy of this data based on the process that exists by default for the static information.

All the **business analysts** suggested that Excel spreadsheets are the tools that are used for all the static data. Nothing sophisticated and nothing complicated. Each data domain owner is responsible for maintaining this master file and once it is validated and confirmed, it will be distributed and uploaded by IT.

Cost against ROI

Implementing an MDM solution is a very complex and costly exercise in which will give many benefits once the process reaches a specific maturity level. Both **CIO**s suggested that the cost that is involved around an MDM process are massive because it requires CAPEX approval for every new software or hardware that is required to be purchased, OPEX approval for all the people that are going to be involved and dedicate time on this process. There must be a sequence of activities for any BTP and this sequence of activities need to start with a centralised MDM solution. "What business does not understand though is the fact that work that is happening in other future projects with regards to DTO would have been eliminated if an MDM process was in place".

The **CFO**s' answers were on the same page with the CIOs' but slightly different. All three suggested that it will be a very expensive process at the beginning with regards to CAPEX and OPEX approvals and justifications, but they think that the most expensive part will be the change of attitude within the organisations. People were used to doing tasks in a specific way and this transition will have to be as "smooth" as possible. What scares them more is ensuring the business continuity and the adoption of the new process from the end-users.

The approach that the **MDM Director** took in this question was different from the previous participants. He focused on the development of an MDM solution as an in-house development or an "off the shelf" product. He suggested that implementing the solution would cost quite a lot of the budget to be spent on hardware, software, and development (internal resources). Also, changes on the existing systems should have to apply the new constraints which are driven by the data governance policy as part of the MDM solution. However, he suggested that with a "proper MDM tool" in place, the time that his team spent on manually managing the content of the data would be reduced. This would enable the team to focus on more productive tasks but also, he could reduce the number of team members. The reason behind this statement was that the checking would now be happening automatically instead of manually. He suggested that with an MDM tool in the place he could reduce the team force by at least 50%.

The **Heads of Development** suggested that the cost that is involved in an MDM implementation is massive for their department and they think that the main cost would be deducted by their budget. They suggested that their team will have to be divided into three

teams. One team to take the responsibility to deliver all the task that are related with the implementation, another team to work on the BAU processes and an extra team dedicated to projects outside the BAU and the Master Data Management. Once the Master Data Management will reach a specific level, all these three teams will have to work together to manage the changes on the current systems which will integrate with the Master Data Management. Then separate testing will have to be involved between these three teams. The business will have to ensure that the integration between the Master Data Management and the existing systems is working as it should work.

The **Program Directors of Business Transformation** suggested that the cost would be huge compared to the cost of their current project. An MDM implementation would involve all this analysis that has already be done to happen again to include the MDM process and the data governance. However, all three were confident that the investment would not add any value once the new systems are in place due to the capability of these systems.

Both **IT Project Managers** have suggested that the cost of implementing an MDM solution would be massive based on the size of the organisation and based on how many functions of the organisation are going to get involved. Another parameter that one of them added was the geography of the organisation. Based on the size of the organisation, distinct functions could be in different geographical locations. This would add more cost to the project's CAPEX and OPEX. They both understood the level of benefits that a process like that would add to the business since objectives like data cleansing, could be removed from the objectives list on a new project.

From a business analysis perspective, the **Head of Business Analysis** suggested that the cost would be quite big since discussions about processes and training (for these processes) should take place before any implementation. They understand that the return on investment would be great, but they suggested (based on experience) that the business does not usually welcome a new "way of doing things" particularly when an updated version that produces the same output is introduced.

Based on the current experience and the time that it takes for an organisation to decide, the **Business Analysts** suggested that the cost would be significant. They suggested that it takes more time for the organisation to decide than move forward with the implementation of any project.

Especially when a project involves both IT and Business and suggestively changes the behaviour of the business.

Time consumption for results

The answers to this question were different between the two **CIO**s. The first CIO suggested that every time he asks for a report from the IT department, he usually gets it within a couple of hours. However, when other functions of the business require reports, IT tries to have an SLA of 1 week before the delivery to the function that requested this report. Once the report is created then everything is refreshable. The second CIO suggested that he never thought it that way. Usually, the problem lays in the lack of a data dictionary. It takes more time to create a data dictionary that has an explanation of the content of the data that the users usually request on a report than doing the report. The problem usually is a result of people deciding new rules on the process, in which they do not inform IT, and when IT processes the data to generate the reports the data do not match.

All the **CFO**s are not aware of this level of detail. However, they all have weekly, monthly, quarterly, and yearly reports. They know that each week the people preparing the report, they start 3 days before the requested day. 1 week for the monthly report, ongoing amendments on the quarterly report and about a month of reconciling the monthly reports with the yearly report because most of the time these reports do not give the same insights.

The **MDM Director** could not give a number with regards to this question. His team constantly receives data or directions from different functions of the team to manage the content and their job is to review the data that they receive every day.

Heads of Development suggested that the most difficult part is to understand where the business needs the data from, how the data link together, and what the content means. One of them suggested: “there are so many different statuses for example for a stock movement that when we want to do a reconciliation of the inventory report, we have new statuses every time, as a result, to go back to the business to ask about this change and then wait for them to come back to us. By the time that they come back to us, we have moved to other BUA requirements

that have come as a priority. Then we go back to the report we take into consideration the recent changes that we had asked for and when we amend these changes there is another status that appeared." (Appendix 4 Question 13)

The **Program Directors of Business Transformation** were not aware of the specific metric, however, they all suggested that with the new systems the reach for information should be very easy. When a question was asked about how much they would think that it would take for a user to find what they need, they suggested that "The user should follow the process that is defined on the system. If the users use the system, the way that they are supposed to use it then they shouldn't have a problem." (Appendix 4 Question 13)

Both **IT Project Managers** described the search for information as "a nightmare". They are not aware of how much time it takes to find the information that they need but both suggested that creating a data dictionary is the most difficult part of any of their projects because each time they have to start from the beginning. One stated that: "In one of the projects we had as a task to create a data dictionary. The project lifecycle was 9 months and the attributes data dictionary took 3 to 6 months and the content data dictionary took 6 months. Even after the delivery of the application, the data dictionary was still not complete." (Appendix 4 Question 13)

The **Head of business analysis** was not aware of these metrics either. His focus was more on the definition of the data content with regards to static information and he described it as a very difficult task. Even if the content of the data is generated within their business function, most of the time, the users know only 40% to 60%. This is challenging for his team of business analysts because most of the times they exclude from the process-specific parameters because nobody knows what they are used for.

Following the answer of the Head of Business Analysis, the **Business Analysts** described the process of identifying the data a very difficult and stressful process since it is often people are not aware of specific content that is very important for the process. They have to deal with altering systems that they are really old and the definition of the "Statuses for example in an order" is not clear. "Users are aware of 5 statuses but when the IT comes back with 28 statuses then there is a problem". (Appendix 4 Question 13)

Cost of time spent searching

Both **CIO**s answered this question in a very simple way. They suggested that every amount of time that is saved on searching data would be a massive benefit for the organisation. Especially when this is a process that should not take time in the first place. However, they suggested that each organisation is different and faces its own challenges. The information, no matter the nature of the organisation should be available instantly and the people who spent more time than the required, they should learn to do their job more efficiently. The other one suggested: "There are some cases that people do the same job for 15 years. This is the only way that they know, these people's tasks are aware only to specific people within the organisation, as a result, IT not being aware of the exact process. When these people are doing something for 15 years, and especially in a dated technology, it is their fault that they have not shared or asked IT of what they are doing. These occasions would probably create false metrics". (Appendix 4 Question 14)

On this question, **CFO**s came up with a more strategic answer. They suggested that nine weeks is nearly a quarter including holidays and absence of the personnel. That means that the best in class organisations regarding data management, they spent almost a quarter of year less than the other organisation on searching. This means that all the other organisations including theirs spent more. This metric triggered question on CFOs like "How much does it cost in our organisation the search for information and how much does it take to produce the information?" (Appendix 4 Question 14) They also suggested that if their organisation spent twice the time searching for the information. It means that their insights are delayed. With that in mind, any decisions that they should have made in a specific time, they are potentially invalid by the time that they get the information.

The **MDM Director** suggested that this number does not come as a surprise to him. Especially, if he compares this number with the time that his team take to go through the content of all the data that they manage. The money that the business could have saved is enormous. First, would have eliminated the overtime and weekend work due to deadlines that the business must meet with customers that receive data from the organisation. Also, any penalties for incorrect data from the customers would be eliminated as well. The other problem that is

depended on specific timeframes is the rebates calculations. When the calculations are wrong and not in time, it means that the supplier will not pay the rebates. This would have a massive impact on the business since the rebates are a large part of their revenue.

The **Heads of Development** suggested by taking this time back, the team of the developers would be able to be assigned to other projects. This time would speed up the process of systems delivery because they (the developers) would have one single source system that they could trust to take the data from.

The **Business Transformation Program Directors** suggested that this would help their project delivery times as the biggest difficulty that they have at the moment is the data mapping and the data definitions which take most of the time. These 9 weeks could be used in different tasks reducing the cost of the BTP and also delivering more accurate and meaningful testing.

The **IT Project Managers** agreeing with the other participants suggested that if they could reduce the time that it takes to do the data mapping and the data definitions as part of each project, then the percentage of success rate would be higher in comparison with the current one. All the projects involve data within IT and usually, the definition is always different and where to find this data is always difficult.

The **Head of Business Analysis** answered this question with reservations. He stated that if the relevant people had a better understanding of their everyday tasks, it would have help other projects with the process analysis. Understanding the information that users are dealing with, it would have speed things up and the definitions of the process would be easier and more accurate.

The **Business Analysts** suggested nearly the same thing as the Head of Business Analysis. They suggested that they spent too much time understanding what the data means for the users that they create it and uses it daily then go there and discuss the process with them.

Mapping Data Sources

Both **CIO**s suggested that each data domain should be maintained in a master source system. The customer usually should be mastered within a CRM, the products within a PIMS, the suppliers

within the ERP along with the employee's domain, the warehouse and the stock location should be maintained within a WMS. The problem starts with the lack of discipline on the systems where customers can be created on the ERPs (when the business has more than one ERP systems), e-commerce platform, CRM, accounting software (potentially), EPOS systems etc. The product can be created everywhere and the supplier on the ERP, the CRM and PIM system. Usually, the actions that are in place are not taken under consideration as the end-users will do whatever it takes to do their job as fast as possible, as a result, this lack of governance and constraints, allows these problems to exist.

The **CFO**s suggested that the ERP should be the main platform that manages the creation of new data and that all the other system should take data from the ERP. Specific governance should be applied to the ERP processes to limit the cases that people create data everywhere.

The **MDM Director** suggested that each data domain should have a main master source system that manages the domain. Once the data is managed in one agreed location, the governance and the maintenance should be easier and the possibility for an error is reduced.

The **Heads of Development** suggested that the main problem that exists is the creation of data all over the data ecosystem. The end-users are not educated, and they are not willing to follow specific directions on how to add and create new data. What the IT uses as the main source is the data warehouse in which potential duplicates can be identified within the dimension tables. Even if this method is not correct or it is not suggested, the people will always try to find ways to do reduce the list of tasks that they have daily by finding a parallel process that is not supposed to allow data creation.

The **Business Transformation Programme Directors** suggested that the ERP should be the main system that is used as the source of data creation. All the processes are linked back to the ERP and from the ERP start all the processes that trigger transactions. Other systems create records, but the ERP should validate these records before processing them. For example, the e-commerce platform or the EPOS systems. These two systems create transactions with most of these transactions to be associated with a new customer. These new records of customers are integrated with ERP. At that stage, ERP should validate the creation of this new record and link it back to the e-commerce platform with a reference to the record that got created in the ERP.

The most frequent problem according to the **IT Project Managers** with regards to the data sources and how to identify which source should be the most important, is the fact that all the sources that allow CRUD are all clustered by their users as important. It takes time to change a process that a business function is using for a long time and has been established as the preferred method for completion. When a new process requires discipline on the way that data is inserted or edited, it usually causes arguments between departments.

The Head of **Business Analysis** suggested that the ERP should be the main system that controls data. All the governance should be applied there and since this is the most important system within the organisation and everything depends on this, there should be more rules applied to the processes.

The **Business Analysts** suggested that the ERP should be the main master source that holds the most accurate information and that all the information that is distributed to any other systems should be the one that comes from the ERP. Also, during the process definition, there have been many data dictionaries that have mapped all the necessary attributes for a process that involves more than one system. This mapping should start from the ERP and should end back to the ERP.

Difficulties during an implementation

Both **CIO**s suggested that all these problems that are mentioned on the question are problems that they have come across during the implementation of every data-related project. There is always the problem that it is described in question 15, where data are stored in too many different systems and every system lacks governance on creation edit and delete of the data. The time is crucial since every day has a cost depending on the number of people that are associated with each project. It also includes the time that is spent before any decisions for starting a project. Lastly, manual data management on Excel spreadsheets. There are other aspects, like agreements between the functions of the organisation that this project is required and why. The rise of CAPEX for each project and how this CAPEX is going to be converted later into an OPEX, how the budgeting will be affected for each year within the IT department for managing the

deliverables of the project and so many other things that cause difficulties before the project start.

The **CFO**s' main consideration was around time and cost. The difficulties that the CFOs' usually have is the justification of the initial cost and the budget allocation. Also, the costs of hardware, software and project time involving employees and external contractors is a challenge. Usually, on any initial project approval discussion, the costs are not accurate. As a result, the new budget should be allocated to the project for completion.

MDM Director suggested that all the above are difficulties for any data related project. To achieve his vision of implementing his data strategy and his MDM process, dedicated members of his team should be focused on specific tasks. Also, the fact that he has to repeatedly prove to the board of directors the value of the MDM and the benefits that a project like this would give the business is very difficult. One of the most important reasons that the MDM Director believes that the implementation of an MDM solution is difficult to be achieved is that "Master Data is too technical for non-technical people, and too simple for technical people" (Appendix 4 question 15). To understand what is required to be done and how. The MDM Director suggested that he spends more time trying to convince the senior management that they need to invest in a solution than planning the strategy for what his role is supposed to deliver.

Both **Heads of Development** suggested that all the above are the main difficulties that they face during any data related project. However, the most important difficulty that they face is the analysis of the process that each project needs to deliver. They suggested that it takes more time to understand what needs to be done and to agree that the final process is accepted between the related functions than delivering the project. Also, the time that is spent on mapping and defining data is a challenge.

The **Business Transformation Program Directors** suggested that the most difficult tasks on any their project is time due to the nature of the business which is heavily based on time constraints and the initial data collection, definition, and mapping in a manual way. However, they all suggested that time is the most critical parameter.

Both **IT Project Managers** suggested that a combination of all four difficulties that is suggested on the question are the most common and from their perspective time is the most

difficult constraint of the project to be achieved. On a project that involves multiple data silos, multiple business functions and multiple different processes, it is very difficult to bring anyone "under the same page" and complete the objectives in the time that is required.

The **Head of business analysis** suggested that on his current project, he faces all the difficulties that are mentioned on the question. Trying to map multiple data between multiple different systems where the knowledge for some of the systems is limited it can be very difficult. time-consuming and with many risks. Then the mappings between the systems are maintained in multiple large and difficult to review and navigate Excel spreadsheets. These actions are part of the initial analysis. Concluding, he suggested that before the project even starts the development phase, the costs are already high.

The **Business Analysts** similarly answered this question with the Head of Business Analysis. They suggested that all these 4 difficulties are visible on their current project and that the most difficult constraint is the time. Having to speak with multiple business users, is proven to be very difficult as they usually have to reschedule process related meetings since the business users do not turn up.

Off the shelf versus In House developed solution

Both **CIO**s, with no hesitation, suggested that they would buy an off the shelf product instead of building an in house one. They both understand that there must be a lot of work that will need to be done before the implementation of the project. When the project and the environment reach the level that is ready for MDM, there will be more potential to deliver the actual value of MDM than building an MDM solution in house. One of them stated, "There is no reason to re-invent the wheel" (Appendix 4 Question 16). All the potential vendors that the CIO would go for, would pass through a due-diligence process. Introductory meetings will have to take place and based on the outcomes of these meetings would decide how they would move forward.

The **CFO**s said that they would prefer an off the shelf solution as well. The reason for that is that they can manage the deliveries better when the conversations and the deliverables are managed within a contract with specific breakpoints. The in-house development usually gets distracted by business as usual tasks that they have to deal with. Also, one of the CFOs stated

that for an important project like MDM, it is always good to have an external vendor to deal with the process. The level of communication between the in-house team and the external team starts from the beginning without any previous history. (Unfortunately, most of the times) In the in-house developed projects, there is a history between departments and individuals that make the project journey more complicated.

The **MDM Director** suggested that he would prefer an in-house MDM solution since he has completed implementation in the past. He understands the process and what needs to be done. Also, it is his strategy that he wants to implement that is why he would not like an external provider to offer a tool that could change his vision.

The **Heads of Development** suggested that an of the shelf solution would be more beneficial for the business since the developers do not have the required experience to move to the next level of detailed and special algorithms. The organisations that the Heads of Development are part of are not software houses. Therefore, the teams have been built up to deliver solutions that will help the organisation to complete specific processes and to support existing ones. Producing sophisticated similarity algorithms and matching mechanisms is not a skill that exists in-house at the moment and there should not be any plan to develop something like that in the future. However, what they are specialised on is the integrations and they would help a lot the MDM process with all the integrations that would be required.

The **Heads of Business Transformation** suggested that they should go with an off-the-shelf solution without a second thought.

The **Project Managers** suggested that the off-the-shelf solution would be the most appropriate. The vendor of an MDM solution has dedicated a research and development department focused specifically on an industrialised solution that works. The IT department with a massive list of business as usual tasks to do daily cannot become research and development overnight.

The **Head of Business Analysis** suggested that he would go with an off-the-shelf solution. He suggested that the vendor will have specific scripts to follow for the implementation as well as a specific set of tasks in a specific order. The Business Analysts in his team would not have the

required experience in defining the Master Data Management process. That is why any exposure would be a benefit for his team.

The **Business Analysts** suggested that the off the shelf solution would be the most appropriate.

Risk of In-House solution

The **CIO**s answered this question by referring to the previous question (18). They understand the risks, and they are not prepared to take them.

On a similar path was the answers from the **CFO**s. They are not prepared to take any risks.

The **MDM Director** was the only one that suggested that he would prefer an in-house developed MDM. He supported that each organisation is different and that needs differ from industry to industry. He suggested that he is aware of the risks that an in-house developed solution would involve. However, he supported that it is more important for him to deliver a solution based on his strategy and not a generic strategy that has been made without taking into consideration the challenges that the organisation faces. With regards to the lack of experience within the IT Development team, he suggested that he would bring in the team external resources to support and supplement the potential gaps that might exist within the development team working together with the CIO and the Head of IT Development.

The **Heads of Development** answered the previous question by answering that they are not prepared to take any risks.

Following the previous answers, all the **Business Transformation Directors** would go with an off the shelf solution because it is too risky to go with an in-house developed Master Data Management solution.

The **IT Project Manager** suggested that they prefer to go with an off the shelf solution because it is less risky than developing one in-house solution.

The **Head of Business Analysis** supported his previous answer where he suggested that the off-the-shelf solution is less risk than the in-house developed MDM solution.

The **Business Analysts** suggested that they assume that the risks will be extremely higher developing a solution in house rather than buying an off the shelf solution.

Data Growth

The **CIO**s were more focused on the infrastructure side of data growth. They said that they could not give an answer based on numbers, but they could say that they predict storage growth for each year for multiple systems and including the data warehouses. The annual OPEX budget for the storage is about 5% of the overall annual IT Budget.

CFOs were excluded from this question as they were not the appropriate audience.

The **MDM Director** suggested that annual growth differs per data domain. He suggested that the product domain shows the growth of about 75% year on year. The customer domain shows the growth of about 25% year on year with some existing customers moving to different competition and the supplier domain shows alteration of about 3% year on year.

The **Heads of Development** said that in regular times during the calendar year they meet with the infrastructure team and they discuss storage requirements and based on the discussions they decide what additional storage they will need for the next quarter if they need any. However, the number of records is not something that they can answer on.

The **Business Transformation Directors** suggested that for their projects, the current standard storage prediction is based on a 3-year plan, with soft evaluation every 3 months and with a detailed evaluation on an anniversary. They could not give a specific number of records as an estimation.

The **Project Managers** suggested that during the industrialisation phase of a project deliverable they estimate the storage that is required to support the application as well as to operate the application on an annual basis. They could not predict growth.

The **Head of Business Analysis** could not answer this question.

Business Analysts could not answer this question.

Unstructured Data Management

Both CIOs described the unstructured data as an aspiration but unfortunately, their organisations are not mature enough for this level of MDM. They both said that they are still struggling with the structured data, so thinking about mastering the unstructured data would not be appropriate at this time.

The **CFO**s suggested that they would like to have more insights from the social media, but they think that the organisations (that each one of them works) are not ready yet for that level of MDM.

The **MDM Director** suggested that he would prefer a separate session for the images of the products which is clustered as content and it is managed within the PIMS but apart from that, there is no plan anytime soon to focus on mastering the unstructured data.

Both **Heads of Development** suggested that there is not any plan for applying any unstructured data in an MDM process.

One of the Three **Business Transformation Project Directors** suggested that there is a plan to analyse social media feeds and categorise them, but this is not involving any MDM. The other two Business Transformation Directors suggested that they do not have any plans to start utilising in a sophisticated way any social media feeds and there are not any plans in the future to use any Master Data Management on any unstructured data.

Both **IT Project Managers** suggested that there is not any plan of using unstructured data on any of their projects. And there is no plan to use any unstructured information outside their typical operational systems.

The **Head of Business Analysis** stated: "The business needs to optimise the way that they work with the structured data and then think about dealing with unstructured data" (Appendix 4 Question 20).

The **Business Analysts** said that unstructured data has not been mentioned on any of the processes that have been discussed so far, and that they doubt that there will be any involvement soon.

Synchronous and Asynchronous Integration

The **CIO**s answered this question without giving exact numbers, but what they said was that all the internal systems (systems within the organisation's network) should integrate with ERP and the ERP should integrate with these systems. Also, they mentioned that in any case that external systems are used, the first point of integration should again be the ERP. Their inspiration is to

achieve as much real-time integration as possible, but this is not always possible due to the systems' limitations.

The question had been asked differently to **CFO**s due to its technical nature. The question had focused only on the importance of synchronous or asynchronous integration. The CFOs suggested that "as long as the information is accurate, it doesn't matter if the information is one or two days later" (Appendix 4 Question 21). Both would prefer as much real-time information as possible, but they realise that it is difficult with the current technology that the business uses.

The **MDM Director** suggested that they have "about 25" internal systems that they should integrate. These 25 systems are separated in multiple ERPs, multiple CRMs, multiple WMSs, multiple PIMs, multiple BI platforms and multiple other systems that are used internally and multiple other system integrations that are used externally. Also, due to the nature of the specific business, multiple EDIs have been set up for a specific type of customers because the organisation is a wholesaler. Some parts of the integrations are as much synchronous as possible but most of the integrations are asynchronous and they run in various times during the day.

Both **Heads of Development** suggested that the integrations that have been developed cannot be measured but new integrations are being built every day. Depending on the type of integration, the frequency of "the calls" vary but mainly, most of the integrations are asynchronous.

The **Business Transformation Programme Directors** suggested that everything should be going back to ERP and critical information should be sent out to other systems from the ERP. The integration with the WMS should be synchronous and real-time, however, when someone places an order, the ERP does not receive it instantly. The business intelligence platforms are being fed information overnight.

The **IT Project Managers** said that they could not number all the integrations within the organisation, but they could suggest that they were "far too many". The majority of the integrations are asynchronous, and this is due to the capabilities of the systems.

The **Head of Business Analysis** suggested that there are multiple integrations between multiple systems. A new integration is designed nearly every day and mainly are asynchronous with most of them being on multiple schedules within an hour.

All the **Business Analysts** suggested that the integrations are multiple, both internal and external with most of them running on schedules and not on events. Real-time-synchronous integration is what the business analyst means by real-time events.

Business capabilities and enablers from MDM

Both **CIO**s suggested that their organisations at the moment are between the baseline and the emerging category by putting their organisation on the step of understanding their data environment and measuring the data quality.

All the **CFO**s suggested that at their current stage, they are trying to understand the quality of their data. They understand, and they are aware that there is a problem however, they will need to know the level of this problem.

The **MDM Director** suggested that he is currently trying to understand the current environment as he is quite fresh to the organisation. Based on this analysis he plans to measure the quality of the data during the analysis of the current environment and then to assess the need of the users who deal with master data to go back to the senior management to secure their buy-in.

The **Heads of Development** suggested that they are currently trying to secure the senior management Buy-in while at the same time trying to measure the quality of the data.
What the **Business Transformation Program Directors** were focusing on is the current quality of the data across the organisation as this is the task that affects their DTO process during their data transformation programme.

The **IT Project Managers** suggested that their organisations are in the same positions as the MDM Director. They are measuring the data quality while they are trying to understand their current data environment. At the same time, they are always trying to make sure that the senior manager will support the MDM process and will authorise the budget for the project.

The **Head of Business Analysis** suggested that the current state is between baseline and emerging. His team of Business Analysts are trying to understand the needs of the business users

that they deal with master data, and at the same time to measure the quality of data across the organisation.

The **Business Analysts** suggested that the current state of the organisation is the assessment of the needs of the business users on using master data and how this need is translated into a process.

MDM Implementation steps of improvement

Both **CIO**s suggested that what they need for their organisation is to implement a formal Master Data Management initiative.

The **CFO**s suggested that it is very important to implement a standardised training for users to understand how to comply with a data governance policy in their daily duties. It is more important for users to be educated and follow the process.

The **MDM Director** suggested that the one thing that he needs, but he does not have at the moment is the support of senior management in practice.

The **Heads of Development** suggested that they should invest more in automation as this would eliminate most of the problems.

The **Business Transformation Program Directors** suggested that they would prefer the business to have invested more in automation before the beginning of their BTP as this would save a significant amount of time and money.

The **Head of Business Analysis** suggested that the business should invest more time and effort on training their users and follow the principles that a process requires.

The **Business Analysts** suggested that Training and automation are the two things that the business should invest in.

Chapter 5: Findings and Cross Case evaluation

5.1. Introduction

This section of the chapter presents the results from both Data Audits and the interviews with a cross-case evaluation. The first part consists of a comparison between company one and company two. The second part consists of an assessment of the Interview results and an understanding of how the opinions and answers are divided between the participants and their line of business followed by an analysis on the impact that the answers have on the MDM and the BTP overall.

There is a summary of the highlights on both audits and the interviews, followed by a comparison between the interviews and the audits, and how the State of the art MDM framework aligns with the audits and the interviews.

As a conclusion to the chapter, a presentation of the three artefacts that represent the contribution of this research is going to relate to the findings.

5.2. Comparison between Company One and Company Two

This section combines the outcomes of the two audits from the two different companies. Both conclusions indicated a problem in both companies which is the lack of data strategy and data planning increasing the risk of failure for both organisations and jeopardising the investments that had already been made.

Both companies were in the direction of using their data as an asset, for that reason there was a transformation process on both organisations. The difference between the two was the fact that Company One was under the impression that there was an MDM process and data governance practice and process. While Company Two, performed the audit since the

management team was already aware that a practice of managing the master data was not in place.

Both companies were hoping to achieve the optimum process on managing their master data and applying a data governance process and both companies had and were willing to increase the budget on achieving this. However, Company One and Company Two were placed in a different timeframe. Company One performed the audit after a failed and costing attempt of a BTP while company two was at the beginning.

Company Two was aware of the situation in comparison to Company One in which there was a misconception of what was needed to happen and what was happening.

In detail, Company One had a team of 19 people lead by an MDM director who was willing to take the responsibility to develop himself an MDM and data governance process. The main problem was that the director was trying to play the role of the business analyst, developer, data architect while he had to perform the daily tasks which were managing the people in his team. At the same time, there was a process that the MDM team had to perform daily which was of high significance to multiple business functions.

The MDM director had documented the MDM roadmap which was included all the eight different steps that are described in chapter 3.4.1. Design areas when planning an MDM function. However, the audit results revealed that the roadmap was only documented. None of the steps that have been defined in chapter 3.4.1. of this thesis had been implemented or the feasibility of implementation of the MDM roadmap was limited.

Company One was facing data quality problems at a high rate. All the three groups defined by Maydanchick (2007) in chapter 3.2. of this thesis were occurring daily and the troubleshooting of these challenges was the daily tasks for the MDM team. The MDM team was also part of the problem as they were creating content and master data as part of their content creation part of their role.

Additional data quality challenges within Company One could be related to the five out the ten different reasons that Sarsfield (2011) defined (chapter 3.2.). The reasons for the data quality problems that Company One was facing and have been defined by Sarsfield (2011) are as follows:

- **Corporate Merges.** (Number four on Sarsfield's list). Company One is one of the biggest companies in their industry, was performing acquisitions of smaller organisations to reduce competition. Due to these merges, new systems were introduced at a high frequency. These systems had to integrate with the existing systems.
- **Metadata Metamorphosis.** (Number eight on Sarsfield's list). Based on the above bullet point, the integrations between existing and new systems required alterations in both systems. These changes had to be incorporated within the data ecosystem.
- **Loss of expertise.** (Number ten on Sarsfield's list). Due to corporate merges, a duplication of personnel is a challenge that most business must overcome. Usually, the resolution to these challenges is the reduction of these duplicate roles. These reductions can lead to a loss of expertise. People that have a system and process knowledge will have to move to different companies and the remaining personnel will have to inherit this knowledge. However, this knowledge transfer is not always complete and results in assumptions. These assumptions could lead to metadata metamorphosis and system alterations. This can be a source of data quality problems.
- **Corporate evolution.** (Number five on Sarsfield's list). Company One and every business will always try to find new sources of income to increase revenue. These newly introduced revenue streams could also lead to metadata metamorphosis. During the audit, Company One was focusing on multiple new revenue streams with the introduction of a new e-commerce platform. This stream required multiple integrations and data transformation from existing systems. At the same time, the existing systems had to be altered to an extension to facilitate these integrations.
- **Information Obfuscation.** (Number two on Sarsfield's list). Being a leader in their industry, Company One had to follow specific regulations in the way that they presented their data. As mentioned in the audit, there are specific bodies that set up the standards for the industry. This forced Company One to receive specific data from these bodies. Company One could not alter this data and should not present or distribute this data in a different format from what they had received. However, there were cases that the controlling bodies

made changes to this format that Company One had to comply with these changes. This action was leading to a different type of metadata metamorphosis that was affecting the data ecosystem, overall.

In addition to all the above, during the audit, there were many instances that the activities and the behaviour from both the business and the MDM team were relating to the bullet points defined on chapter 3.6. and the worst practises in enterprise data governance. However, the most important was the perception of the MDM director who was under the impression that he could deliver an MDM solution including business analysis, design, development, maintain and communicate the governance by himself. The MDM director was trying to convince the business that he could deliver this process on his own. This links to what Woodie (2016) suggested as "Relying too much on Unicorns". One person cannot deliver MDM and data governance on their own. These processes are formed by the principles of cooperation and require the entire business to participate. Otherwise, there will be a situation that Sherman described as "failure to implement" (Sherman 2011). This is usually the non-enforcement of the above in any business process. In the case of Company One, the data governance effort would not produce any business value. The data governance process should be recurring feedback, in which data is defined, monitored, acted upon, and changed when appropriate (Sherman 2011).

Company Two was facing different challenges from Company One. As per the company background, Company Two did not have an MDM and data governance policy in place. It was a greenfield environment for a BTP. Company Two had decided to change the way that they operate and alter the business to be completely data-driven.

There was not an MDM and data governance roadmap, and they needed a role that could lead to this function. However, Company Two had to overcome different challenges. These challenges were more related to political issues and intercompany relationships between the BTP team and the IT team. Communication and cooperation were broken. As a result, every team was working in isolation believing that their approach would be the most beneficial for the business. The architecture of MDM and the roadmap of achieving that was not available but there were instances of bad practices before the initiation of the MDM process.

The results of the audit can be linked to all the ten reasons that Sherman (2011) identified as of substantial risk in an MDM and data governance implementation.

- **Buy-in but not commitment**. Within Company Two the main two project teams work in isolation. Even though both IT and BTP team understand the need for data governance and MDM policy, the do not engage to each other as a result of one team to proceed with tasks that should require involvement from both teams. The data definitions between existing ERP and the new ERP are not aligned and the BTP team engage mainly with the vendor and they do not communicate these definitions to the IT team.
- **Read, fire, aim**. A prime example of Sherman's second bullet point is the fact that the WMS implementation runs in isolation from the ERP project even though both systems should be aligned. As a result, WMS implementation to be based on assumptions that specific data definitions would be available. However, these definitions are not guaranteed and if there is an alteration, there is a possibility that it would not be communicated in time due to both teams operate is silos. The result was that WMS would be ready before ERP and would remain idle until the ERP implementation completed.
- **Not dealing with change management**. The communication problem between the two teams increases the risk.
- **Trying to solve world hunger or boil the ocean and goldilocks syndrome**. These two bullet points from Sherman (2011) can be combined in Company Two. There were instances that both teams were trying to define processes to the lowest level of granularity but at the same time, they were trying to deliver specific schedules in predefined timeframes. As a result, both teams were spending more time in defining metadata and there were many instances that there were deliveries with irrelevant results from both teams. Both teams in some instances were acting locally but they were not thinking globally (Sherman 2011) and they were trying to be "Too big, too fast" (Sherman 2011).
- **Technology alone is the answer, ignoring data shadow systems, not having a sustainable ongoing process and failure to implement**. These four bullet points from Sherman (2011) could represent the situation within Company Two. BTP team was supporting the fact that

the new ERP would be the solution to every challenge that Company Two had. However, the BTP team was mainly focused on the new ERP implementation and they were ignoring any other systems that the IT team knew. As a result, the BTP team was building on sustainable processes ignoring these systems. When the testing of these processes was under testing the results were not satisfactory and a revisit on the process design had to be re-occurred.

Overall, the main reason for the challenges that both companies were facing, was the communication between business functions and the competitive spirit between the teams. Both Companies needed a role that could take lead on both programmes and communicate accordingly. This role had to be independent and bypass any political frictions between the organisations' functions and teams and approach the implementation of an MDM and data governance process in a neutral way.

The table below summarises the key findings from the audits.

Highlights	Company One	Company Two
Position in BTP	Company One performed the audit after a failed and costing attempt of a BTP	Company Two performed the audit at the beginning of a BTP
Data State Awareness	Company One was under the impression that there was an MDM process and data governance practice and process	Company Two, performed the audit since the management team was already aware that a practice of managing the master data was not in place.
Data Quality Challenges	•Corporate Merges. •Metadata Metamorphosis. •Loss of expertise. •Corporate evolution. •Information Obfuscation.	Greenfield environment. Company two was aware of the existing Data Quality challenges, that is why they wanted to invest in Transformation to overcome that.
Data Strategy	Data Strategy in Theory but not in practice.	Company Two was aware that there was not a Data Strategy in place, that is why they performed a data audit.

Use Data as an Asset	Company One was highly depended on their Data, however, there were many challenges on using the data as an asset. The main use of data was to fulfil operational needs on a request basis.	Company Two's goal was to transform completely into a data-driven organisation
MDM process	Falsely under the impression that there was an MDM process in place. There was a team of 19 people under an MDM team.	Fully aware that there was not an MDM process in place.
Data Governance	Trying to apply some form of Data Governance	Fully aware that there was not a Data Governance practice in place.
Hope on achieving MDM	Hope to achieve in the future	Hope to achieve in the future
Hope on applying Data Governance	Hope to achieve in the future	Hope to achieve in the future
Willing to increase Budget	Willing to increase the budget to achieve better Data Quality	Willing to increase the budget to transform the business into a data-driven organisation
Believe in Unicorns	MDM director was trying to take the role of MDM architect, developer, data analyst, process analyst, specification writer etc.	Fully aware of the limitations.
Change Management	Company One had change management problems due to data Quality reasons.	Communication problems between teams.
Technology is the answer	Falsely under the impression that technology would solve all the problems on its own.	To an extend there was a perception that technology would eradicate all the data quality problems of the past.
Intercompany Relationships	Highly competitive.	Highly political
Communication	Trying to convince Senior Management.	Teams working in silos and not updating the Programme's progress.

Table 14. Key Findings from Data Audits.

5.3. Critical Reflection of the Interviews

5.3.1. Data Quality and Master Data, Question 3 to 5

5.3.1.1. Agreements and disagreements between the groups.

All the participants agreed that data quality is important and vital for their projects and the business in general. Each one of them for their reasons. However, in most cases, the project managers and the project directors were prioritising the delivery of their respective projects based on the deliverables. To some extent, meeting the delivery schedule was more important than the data transition from the old systems to the new. The business analysts were more focused on analysing and designing existing and future processes assuming that the data quality should be a concern for a different department.

Regarding the definition of master data, all the participant used words including “clean”, “truth”, and “golden record”, “de-duped”, “static data”.

5.3.1.2. Impact on MDM and BTP

Based on the results. At that early stage of the interview, there was an understanding that the participants in their projects were following Joshi’s methodology as described in chapter 3.4.1. The two points on Joshi’s methodology that was described from the participants were the points 3 and 7, “Collecting business metadata” (Joshi 2007) and “Collection and maintaining metadata between technical and business rules” (Joshi 2007). However, there were two risks which match with Sherman’s approach. All the participants were focusing only on their BTP and the new ERP and they were ignoring any other systems and the integration between the existing environment and the upcoming. There was also an assumption that the new ERP system would be able to manage all the master data as part of the default functionality. As a result, there was not any focus on the data governance which according to Woodie (2011) is a very substantial risk.

5.3.2. Master Data Management Awareness Questions 6 to 8

5.3.2.1. Agreements and disagreements between the groups

All the participants agreed that managing their master data is very important for all their projects and their businesses in general. However, every group of participants had a different focus. For the CIOs, the focus was the governance regarding the master data and how this would affect their function. The CFOs were more interested in the accuracy of their numbers and they were very positive in supporting the overall data management to get the accuracy that they seek. The MDM director was focused on data governance as a policy and described the MDM as a business as the usual process for his department. The Heads of Development were focused on the data integration and assuming that the data is accurate, all their integrations would have a better success rate.

However, the project teams including the Business Transformation Program Directors, the Project Managers and the Business Analysis teams were more focused on the delivery and the definition of the processes rather than the data. Their most important concern was the timeframes and the budget, and they would investigate further the master data process only if there was part of their project plan. However, the Business Transformation Program Directors had a very different view of the data. One of them suggested that if the data on the existing systems is not accurate, they would not spend any time to enhance and increase the validity of their data, but they would just ignore and focus on the new data sets that the new systems would generate. All the members of the different project teams were more focused on managing the associated budgets and mark the processes as completed.

All the participants agreed with the current state of managing the master data, and all of them suggested that there was not a properly set up process for this type of activity. However, the MDM Director was the only one who was suggesting that the process is in place in principle, but there is not a completed overall process across the business. Also, the MDM Director, suggested that there was not the optimum support from the business. Several groups were

working in different parts of the projects dealing with master data, but the overall feeling was that everybody was working in isolation.

The final point from this group of questions suggested that all the participants agreed regarding future expectations. All the participants were positive for the future despite the challenges that they were facing during their projects and after the completion of their projects.

5.3.2.2. Impact on MDM and BTP

Overall, there was a positive understanding of the importance of MDM from most of the participants. The important outcome is that almost every participant understands the challenges that they currently face within their business and their project regarding the MDM in general. And the most positive outcome is that almost everybody optimistic about the future. This optimism suggests that there are good foundations and willingness within the business for future improvements that are beneficial for the future of MDM and the success of an overall business transformation.

However, some elements can be classified as a risk. The highest risk is the fact that the BTP Directors are principally focused on the process and the deliverables but not on the data. The most important concern was from one of the participants suggested that if the data is not in a good state, they would just ignore it. This comment can be very dangerous in any kind of project and not only on a BTP. They bypass completely the data history and put in risk the business continuity.

An additional risk is a fact that MDM Director suggests that there is not enough support from the senior management and that all the teams are isolated when it comes to master data. He believes that there is a data governance process but there is not a system in place. However, he is in discussions with multiple vendors who provide MDM solutions. This is also a risk since based on the conversations, there is not a complete picture of the overall data map and both his team and the project teams work in isolation. That means that there is no coordination and schedule. Based on that, there is an attempt that his efforts may be characterised as invasive in case he tries to enforce a data governance policy.

5.3.3. Data Strategy Questions 9 to 22

5.3.3.1. Agreements and disagreements between the groups

In this section, the participants were called to discuss the data strategy within their current projects and within the organisation, overall. The CIOs' objectives from their data strategy are to achieve data quality, ease of access to the required information and to meet the business objectives. As part of their **data strategy** is also the archiving of their data. The CFOs' objective was accurate reports through consolidated business insights. The MDM Director's objective is to align people with process and technology to achieve data governance. The Heads of Development suggested that the important outcome would be to understand the business rules and to embed these rules into their integration strategy. The Business Transformation Program directors suggested that the data strategy is not their focus as they are interested in ensuring the business processes within the new software. They also believe that the new ERP would have embedded data governance rules. The rest of the participants suggested that their focus was on the business rules and how these rules relate to disparate business processes.

There are different views regarding the **support** that data strategy receives from the **senior management**. CIOs agreed that it is challenging to convince business to support technology since this part of the technology does not have any visible effects on the processes. However, CFOs belief was that they support the data strategy. Business Transformation Directors were aligned with the CFOs'. They believed that they support the data strategy that is why there is an investment in new systems. All the remaining participants, however, they stated that senior management supports MDM but not in practice. They suggested that there is always an agreement of what is required to be done, but the engagement is not always guaranteed.

All the groups agreed that achieving **data consistency** across all the functions and systems of the business was a challenging task, especially when all the static information is managed manually. All the participants suggested that the master data or static data is maintained and amended manually within Excel spreadsheets. The BTP directors suggested that all the manual processes regarding data management will be managed from the new ERP system.

All the participants agreed that an MDM implementation comes with an immense **cost**. Equally, all the participants agreed, however, that the future benefits would balance the original costs. The CIOs, CFOs and the MDM Director supported that their **investment** could improve their operations and reduce the number of employees. The Heads of Development suggested that part of their teams that are dedicated to managing data for multiple projects would return to their original tasks. By achieving that, the development departments would be able to utilise all the members of their teams and focus on their original tasks. All the other participants that are part of the project teams suggested that any future costs would be reduced. The reason for this was based on meetings and discussions over data cleansing and data mapping would not be necessary for future projects or future amendments in business processes.

With regards to the **time that is required to get data related results**. All the participants suggested that each function within an organisation and each role has different objectives. Every function requires a different type of results in a different frequency. However, all the participants suggested that the reason that an answer based on data is often due to lack of knowledge or not enough expertise on business processes. The **time** that is **spent** on **seeking information** through the data has also a significant **cost**. All the participants agreed that a Master Data Management and data governance process would significantly improve the utilisation of the employees. The CFOs also suggested that their strategic decisions would be based on more accurate and up to date results.

To achieve the anticipated return on investment, the **data sources** need to be **mapped.** According to the CIOs, however, the general perspective is that there is a lack of discipline in managing master data. Source systems and master data are created and used in multiple environments within the business very frequently and there is not a specific way that this can be controlled. To avoid this challenge the CFOs and all the participants who are members of a project team, believe that the master source for data creation and maintenance should be only the ERP. The MDM Director and the Heads of development agreed with the CIOs that due to lack of governance and discipline, every source should be controlled. Even though a data governance policy clearly defines a specific way of managing the data, users would always find a way to overpass this policy.

The above also can be characterised as one of the **challenges during an implementation.** According to the CIOs, the vast amount of data sources (especially Excel spreadsheets) and the lack of discipline from the users are the biggest challenges. For the CFOs, the re-evaluation of the budget due to cost and time that each project requires is the biggest challenge. The MDM Director, however, suggests that the biggest challenge that he must overcome is the time that he spends convincing the senior management of applying and enforcing specific rules on data creation and maintenance. He suggests that master data is very technical for non-technical people and very business-oriented for the technical people. All the other participants agreed that the business process knowledge and understanding is not always completed. The missing element, no matter the size, is always the most challenging part. The Heads of Business Analysis and the Project Managers agreed with the CIOs that the challenging part is mapping of all the disparate data sources. The Business Transformation Directors suggested that the most challenging part is the data collection, the definition, and the mapping of the data.

When it comes to the **preferred solution**, all the participants apart from the MDM director, stated that they would strongly suggest as the best option, an existing solution. The MDM director suggested that an in-house is the most preferable option since there is no limitation on implementing his strategy. He would overcome the challenge of lack of subject matter expertise by acquiring short term external resources and subject matter experts. However, all the other participants suggested that this approach would be a **considerable risk**.

With regards to **data growth,** CIOs suggested that 5% of the overall IT budget is dedicated to data growth as an operational expenditure. The MDM Director suggested that the data growth differs between data domains and the Project Managers suggested that the required infrastructure according to their project plan is always estimated based on the initial requirements of the project deliverables.

None of the participants suggested that their focus is **unstructured data.** However, the CFOs and the Business Transformation Directors suggested that they would like in the future to understand the data from their social media feeds. However, the priority was structured data.

The type of integration that all the participants suggested that is used as the **asynchronous**. Every activity was happening based on a schedule, and not due to an event. The CFOs suggested that accuracy is more important than a real-time (**synchronous**) integration.

At the end of the conversation, all the participants suggested that understanding the level of the data quality and measuring the data quality was one of the actions that there were undertaking. However, for a **Master Data Management implementation improvement**, the most suggested activities were the training of the end-users in how to comply with data governance (CFOs, Head of Business Analysis and Business Analysts). The CIOs suggested that they need to implement an MDM initiative. The Heads of Development suggested that they need to improve their automation process. The Business Transformation Program Directors agreed with the Heads, but they suggested the automation should be in place before the initiation of the BTP. The MDM Director suggested that improving an MDM implementation requires the support of senior management.

5.3.3.2. Impact on MDM and BTP

Based on the answers that the participants gave during the conversations, there are a lot of elements that their BTP and their MDM implementation are at risk. All the participants from all the different groups understand the need for data quality. However, there is not an activity in place to ensure that data quality can be ensured. Based on Chapter 3.2, Maydanchick (2007) divided the data quality problems into three distinct categories. The processes that bring data from outside the database, the processes that use data that are handled inside the database, and the data that already exists in the database but becomes obsolete mainly due to time, and not due to changes in the database. The answers from the participants indicate that the current process of managing static information is a significant risk to the BTP and the MDM as a process. The data consistency cannot be guaranteed due to the excessive use of manually maintained Excel spreadsheets. The Excel spreadsheets do not have any control over the input from the user, and typographical mistakes can happen without capture. The fact that these Excel spreadsheets

are used to be inserted into different systems that are designed to provide important business insights or perform important business processes, increases the risk. All the participants also mentioned that the amount of disparate data sources that use the same data cannot be easily consolidated because of multiple constraints including time, complexity, and lack of understanding. All these reasons allow the business processes to operate without any control but also any validation. Also, many participants suggested that there is a lack of understanding with regards to business processes but also a lack of reasoning of why a process needs to follow a specific execution. This suggests that this process may alter data that already exists in the database even though this data should not be changed. Maydanchick's (2007) reasoning for data quality problems is aligning with Sharsfield (2011). On chapter 3.2. Sharsfield's (2011) list or solutions to data quality problems is explained. However, none of the solutions that Sarsfield suggests is followed in practice based on the answers that the participants gave.

MDM should be the strategic solution to the problem according to McKnight (2009). In his definition, McKnight (2009) suggests that MDM itself cannot resolve the data quality issues and that organisational changes by the senior management need to happen to achieve success. Based on the answers that the participants gave, there is a conflict of what senior management is prepared to do, what it is expected to do and is happening. Most of the participants believe that there is not practical support from the senior management and at the same time, senior management believes that the required support is offered. All the participants desire the benefits of quality data, but all of them believe that the responsibility lies within someone else's territory.

The environment that is described can be related to what Sherman (2011) suggested as "Buy-in but not commitment" in Chapter 3.6. The business accepts and understands the reason of why a data governance programme needs to be done, they put the people in the right places but when it comes to them to do their tasks, the engagement is very limited to none. Few of the participants, especially from the BTP team, indicated that there was an activity that should happen before the BTP initiation. These actions are initial data collection, definition, and mapping. This also links to Sherman's (2011) "Ready, Fire, aim" point in which the organisations are trying to assign the people and form panels before the in-depth understanding of what the

scope and the real purpose of the data governance programme, as well as its roles and responsibilities of the participants, is.

Few participants also answered that end-users could not decide or explain which data source is the master, as a result, they were insisting that everything should be used and be taken into consideration during the data definition and data mapping. According to Sherman (2011), they are “trying to solve world hunger or boil the ocean”. The participants suggested that the business believe that it was feasible to solve all data problems in the initial phase of the project. This is a risk and as the CFOs answered in one of the questions, they usually must re-evaluate the budgets for covering the extra costs.

One of the biggest concerns, however, is the selection between an existing solution or an in-house built solution. Every participant suggested that the most appropriate solution would be an “Off the shelf” solution. The MDM director suggested that his preferred option would be to build one himself. The problem in this phrase is that the MDM Director according to Woodie (2016) acts like “a Unicorn”. It is not clear if he is trying to prove himself to the senior management or he truly believes that he can succeed with this task. However, both options are contradicted by his other comments based on the overall conversation. He has mentioned that the most challenging part of the implementation is the time that he spends trying to convince senior management, but also trying to understand from the business processes and rules. According to his sayings, this is the most challenging part. The question, however, is why it would be any different if an in-house MDM solution project is approved by the business. The business users would still operate in the same way as all the other participants have suggested this kind of behaviour by the users.

The main point that it needs to be taken into consideration, is that all these organisations that the participants are engaged with, are not software houses and they do not require a research and development department to deliver new and revolutionary software. The resources that are part of the IT function are there to support a process and increase the revenue for the business. They are retailers and wholesalers that their main objective is to sell their products or services that they make them unique in the industry. Investing in an in-house MDM solution

would be a very expensive exercise to produce a product with no intention to sell it or offer it as a service to a different customer.

5.3.4. Questionnaire Results Questions 23 to 27

The perspective of Master Data Management within the organisations

Based on the answers that the participants gave, there is a trend between the Functions. Business and IT people have a different perspective with regards to the importance of the MDM to core business operations. IT and IT-related participants have a different view of the project team and the business analysts. The surprising fact is that one of the Business Transformation Director believes that the process will resolve any potential issues and that it will not cause any future issues. The same trend applies to the question regarding the importance of buy-in from senior management as well as the resources and budget support for Master Data Management. This trend changes when it comes to adherence to master data policies which everyone seems to highly support the importance. Regarding the current trust in data systems and policies, the participants marked with the same score which indicates that there is not enough trust. However, they all believe that once the new systems will be in place the master data will be accurate and trustworthy. All the interviewees agree that it is very difficult to identify any records with significant errors and nearly all of the support that these errors are very difficult to be fixed. All the interviewees believe that it is quite difficult to gain any information that is required and nearly all of them believe that at the current state a decision based on data would take longer than it should take. With regards to new data integrations, the same trend appears between the IT and the project teams. The IT and IT-related interviewees believe that it is not that difficult while the project team members believe that it is quite difficult. However, the gap between them is not big this time. With regards to time reduction in data-centric processes, all of them have the same belief. This belief indicates that people will still need to spend time on data-centric processes.

The following graphs represent the results from the participants. Due to the variety of the participants, different roles, different points of view within the same roles and also due to the small number of samples, each question has 3 sets of graphical presentations one on average per function, one on average per role and one with the actual score per interviewee. All the answers can be seen in a table format in Appendix 1.

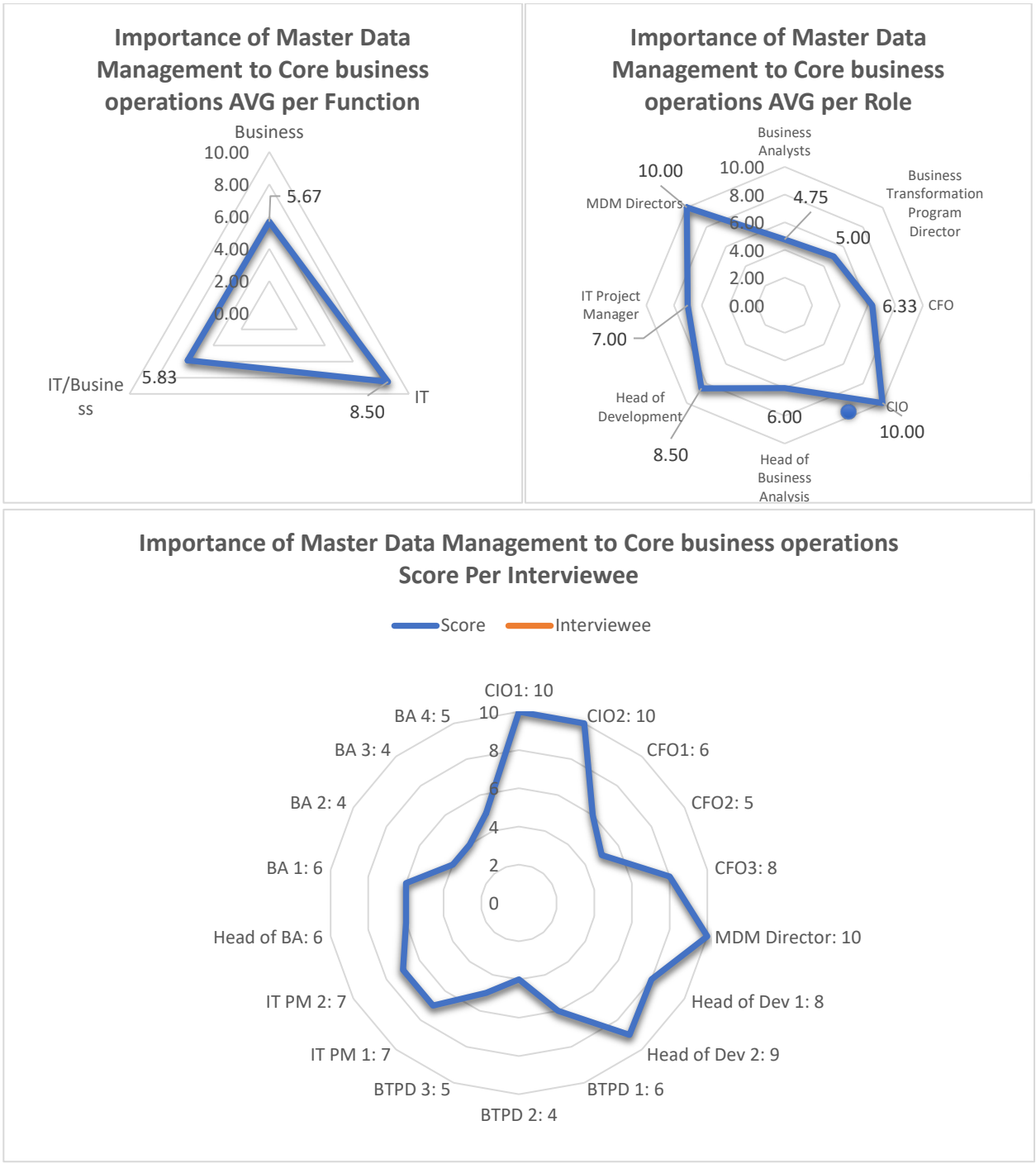


Figure 3 Question 23 - Importance of Master Data Management to Core business operations AVG/Function, AVG/Role, Actual Score/Interviewee

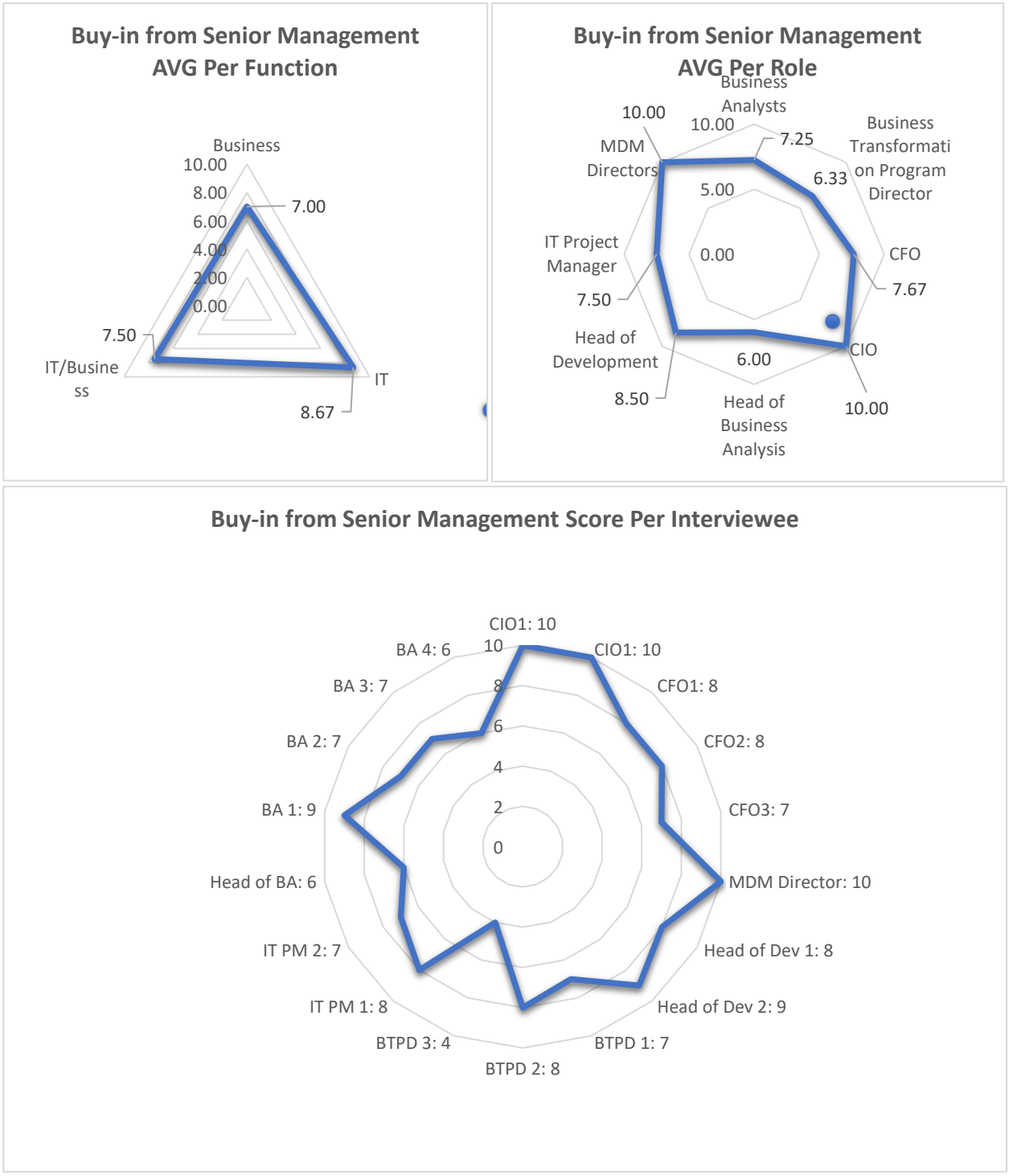


Figure 4 Question 23 - Buy-in from Senior Management AVG/Function, AVG/Role, Actual Score/Interviewee

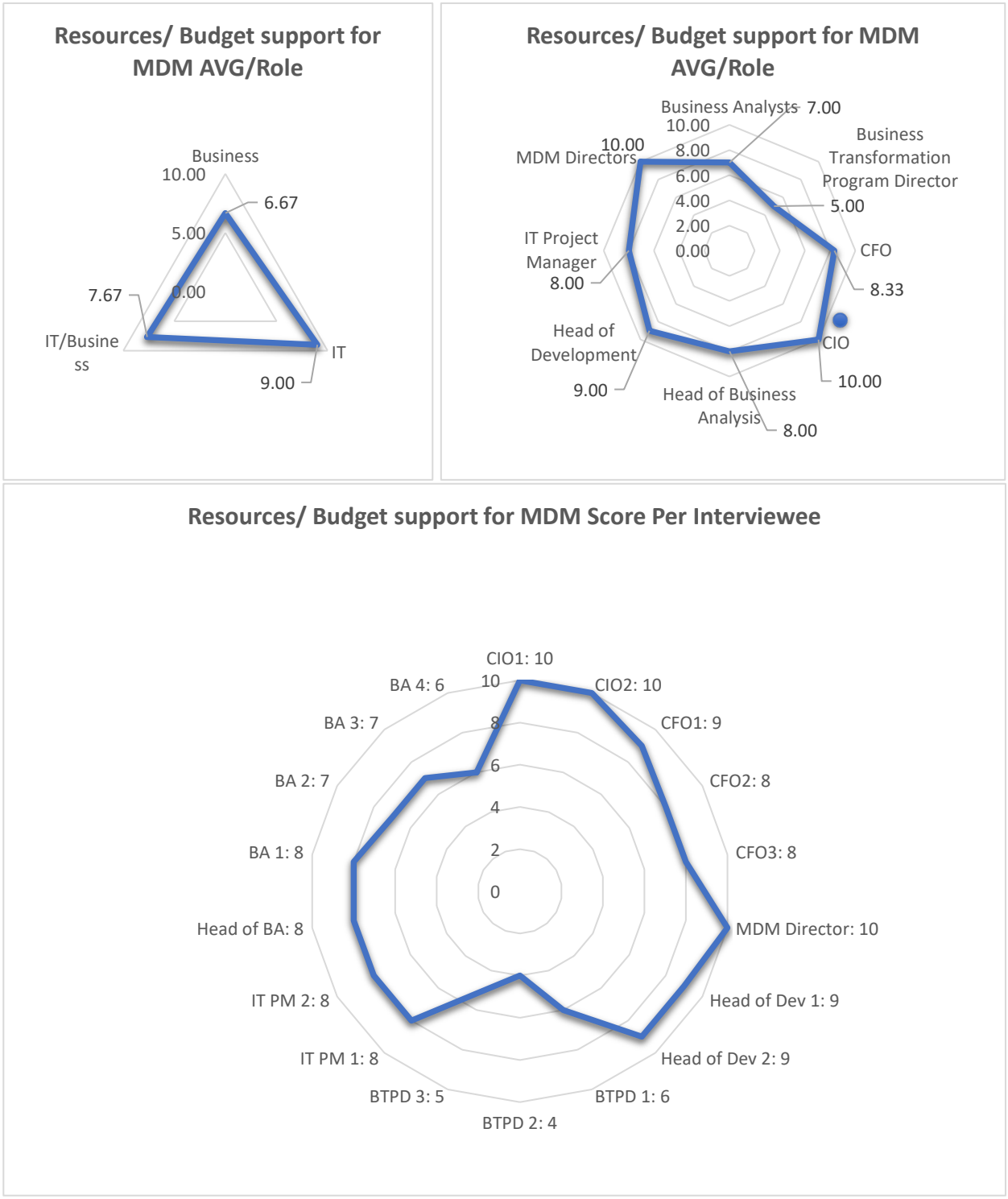


Figure 5 Question 23 - Resources/ Budget support for MDM AVG/Function, AVG/Role, Actual Score/Interviewee

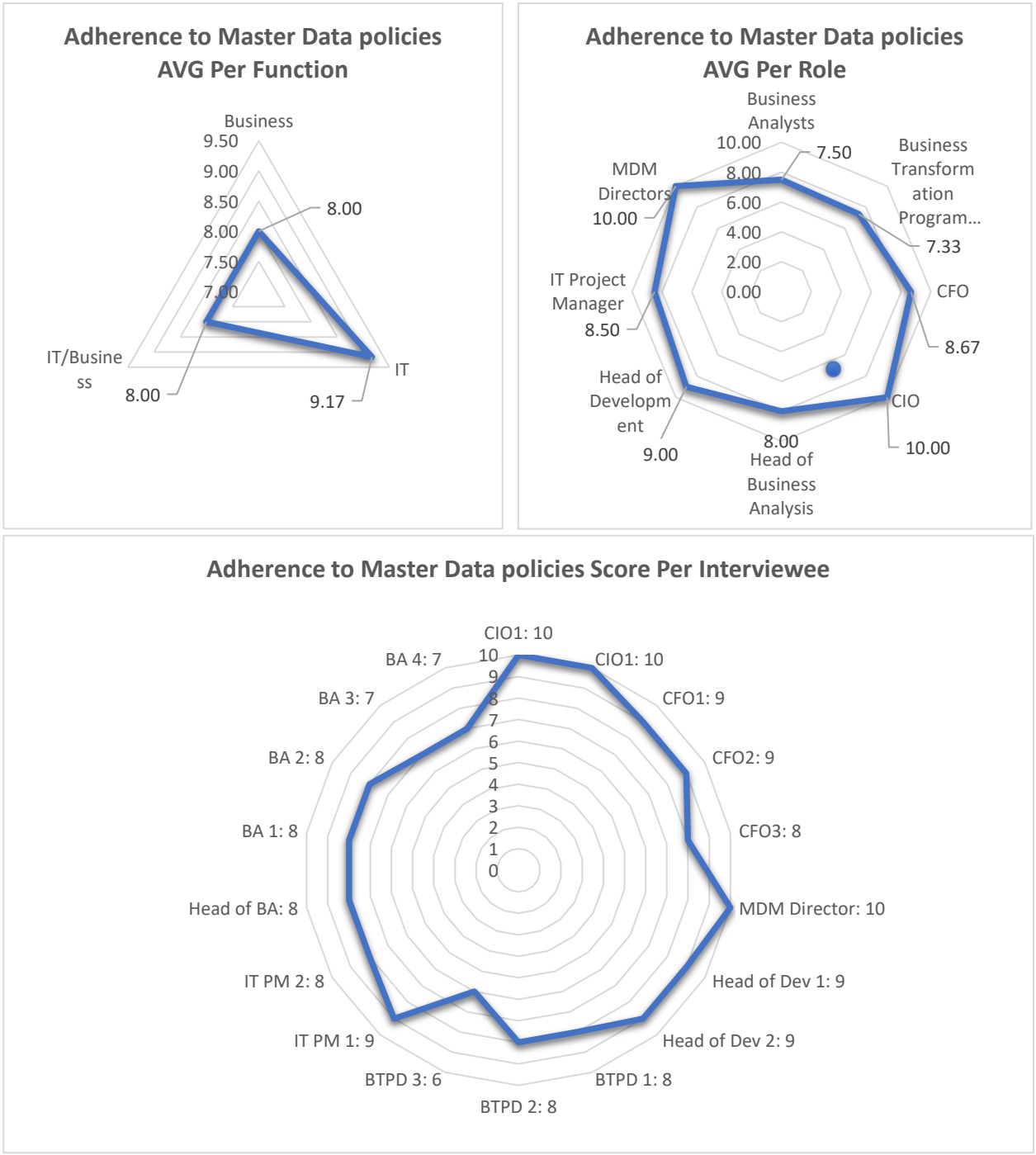


Figure 6 Question 23 - Adherence to Master Data policies AVG/Function, AVG/Role, Actual Score/Interviewee

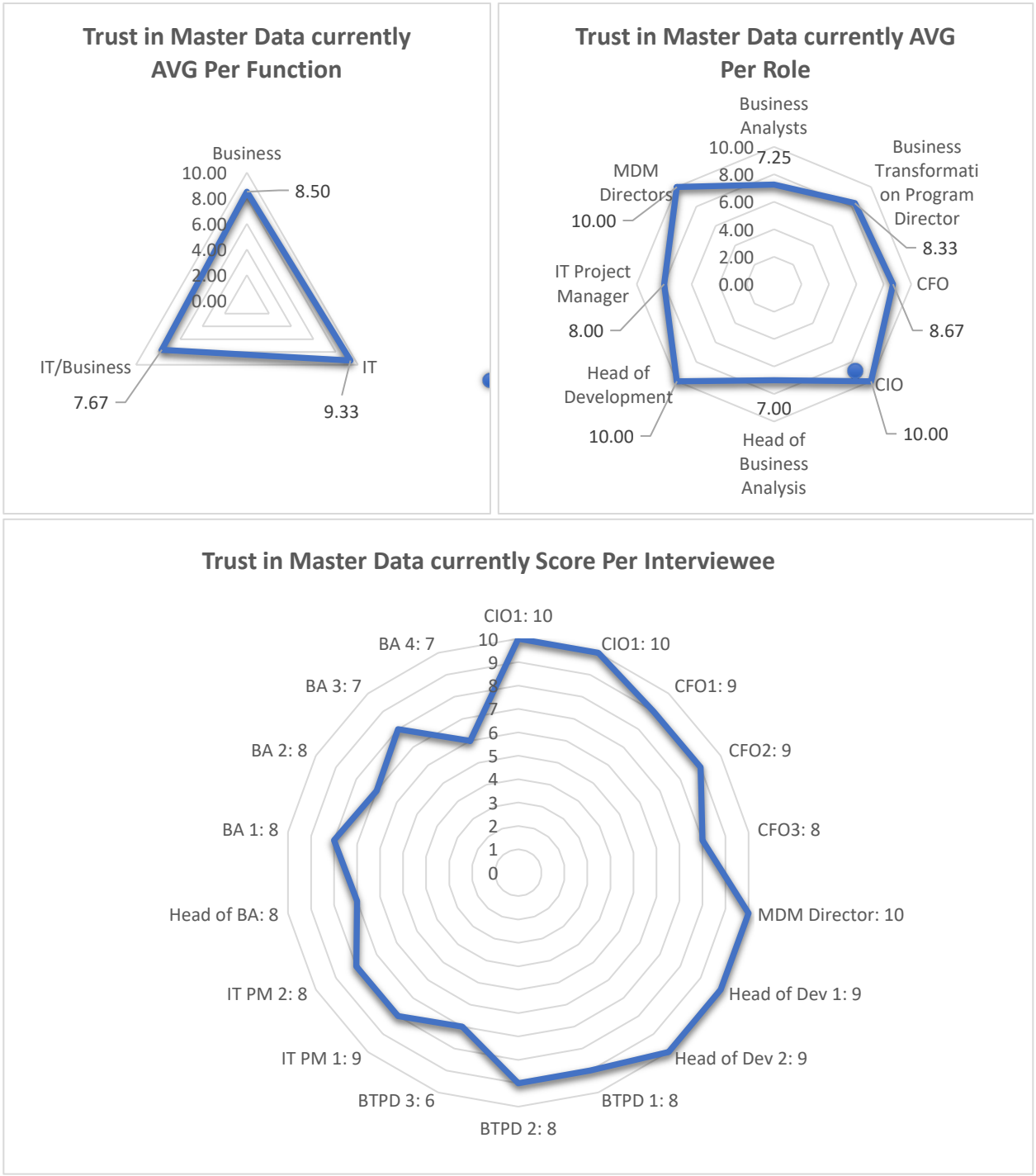


Figure 7 Question 23 - Trust in Master Data currently Score Per Interviewee AVG/Function, AVG/Role, Actual Score/Interviewee

Figure 8 Question 23 - Trust in Data systems and Policies currently Score Per Interviewee AVG/Function, AVG/Role, Actual Score/Interviewee

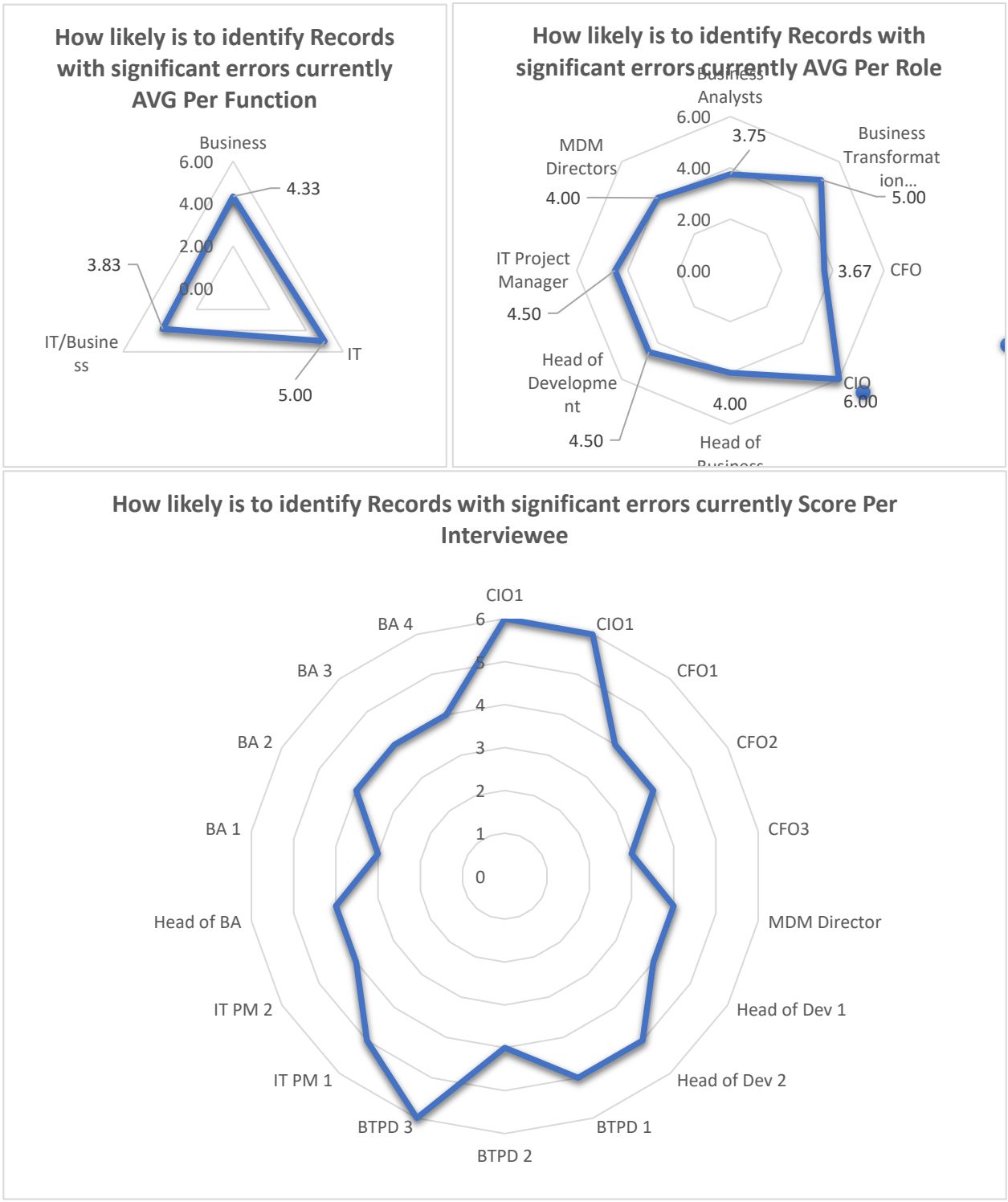


Figure 9 Question 23 - How likely is to identify Records with significant errors currently Per Interviewee AVG/Function, AVG/Role, Actual Score/Interviewee

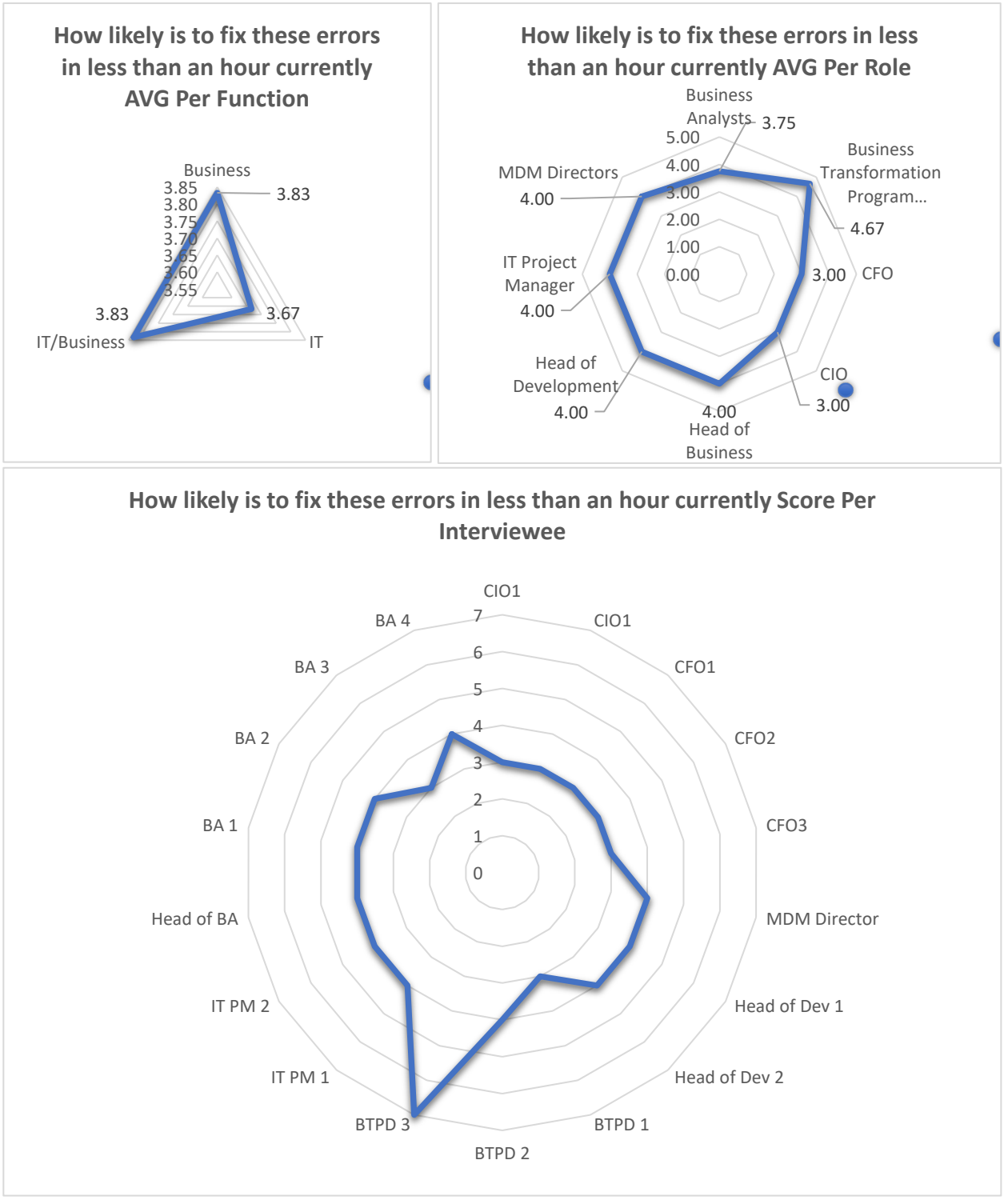


Figure 10 Question 23 - How likely is to fix these errors in less than an hour currently Per Interviewee AVG/Function, AVG/Role, Actual Score/Interviewee

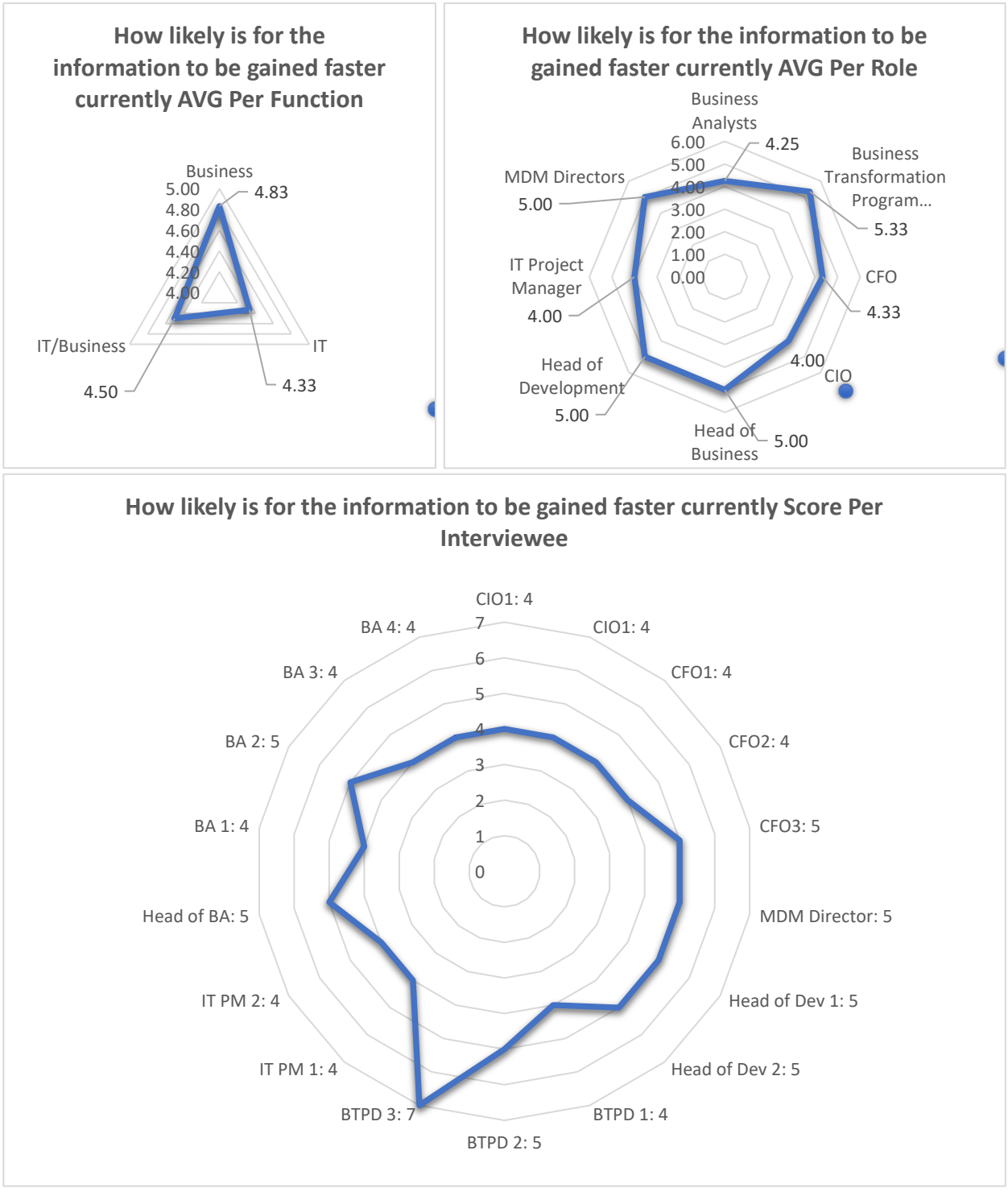


Figure 11 Question 23 - How likely is for the information to be gained faster currently AVG/Function, AVG/Role, Actual Score/Interviewee

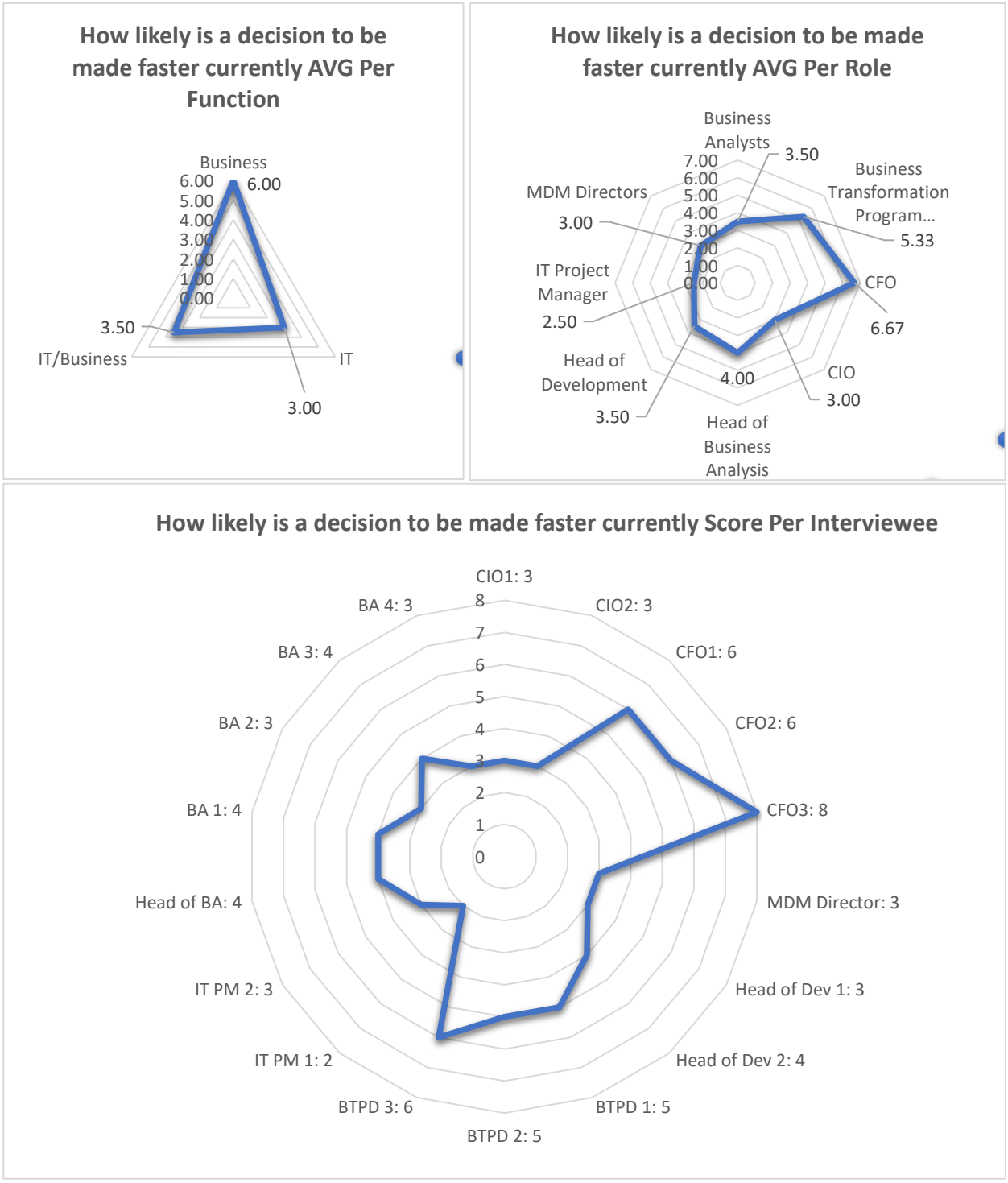


Figure 12 Question 23 - How likely is for the information to be gained faster currently AVG/Function, AVG/Role, Actual Score/Interviewee

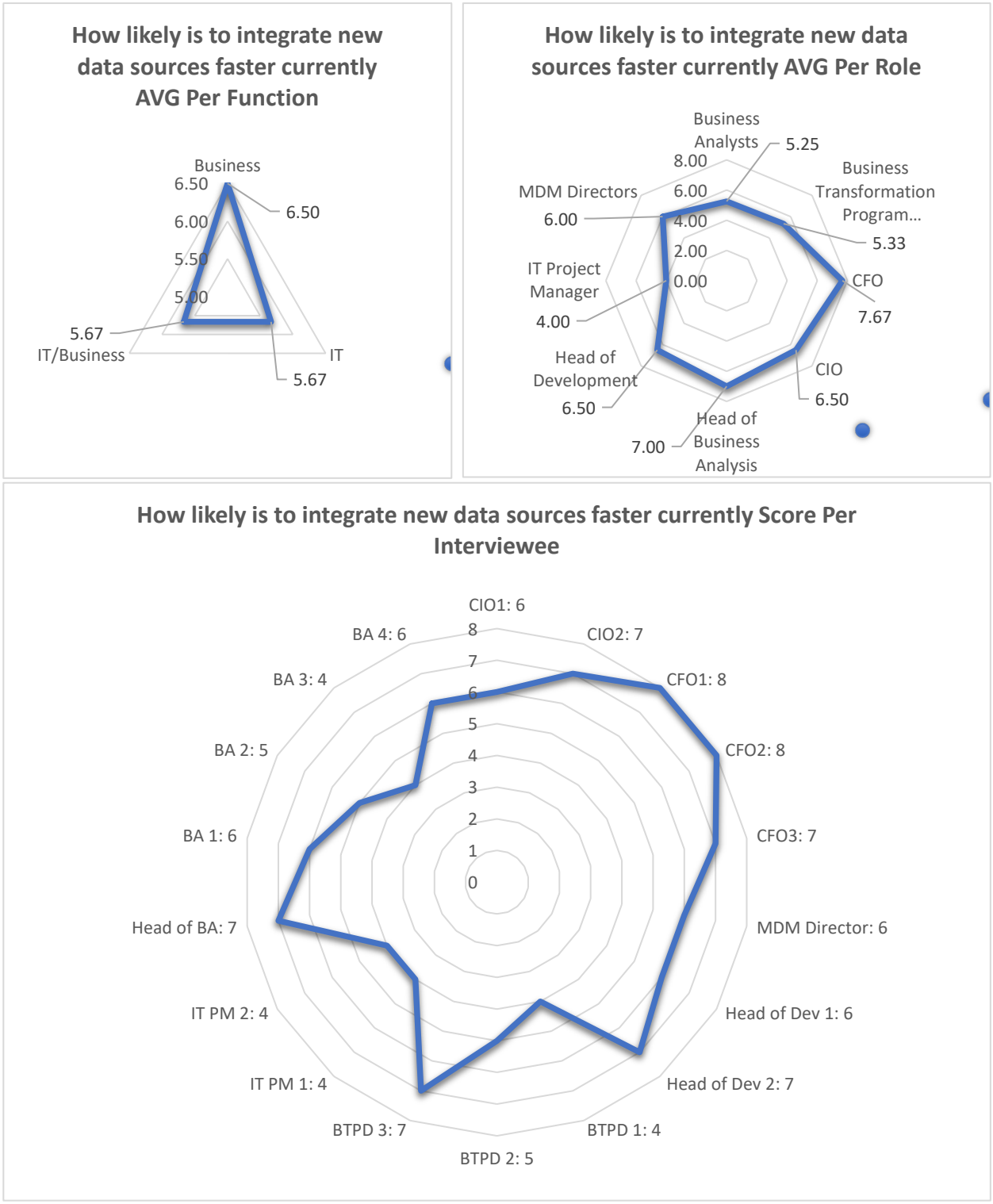


Figure 13 Question 23 - How likely is to integrate new data sources faster currently AVG/Function, AVG/Role, Actual Score/Interviewee

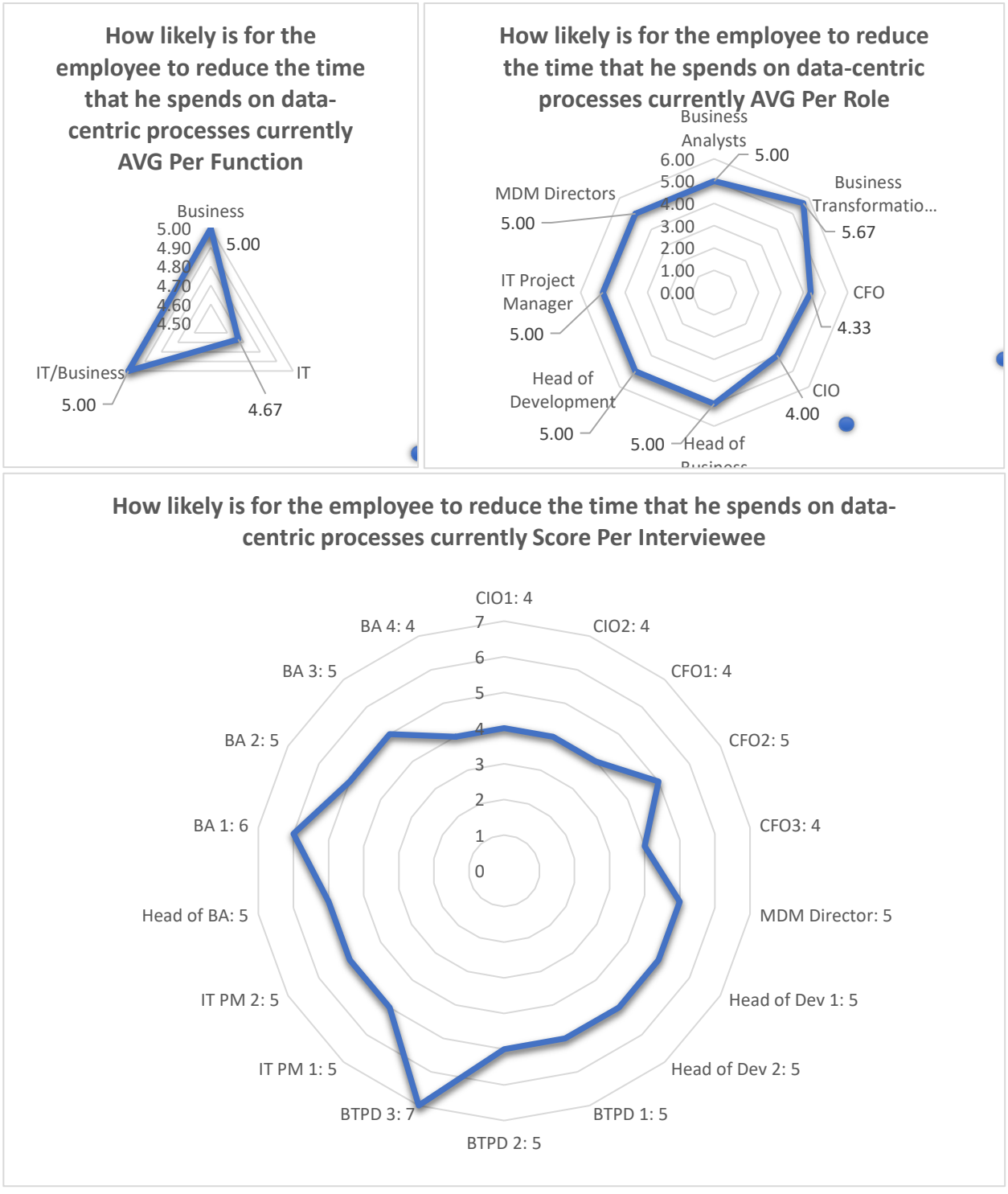


Figure 14 Question 23 - How likely is for the employee to reduce the time that he spends on data-centric processes currently AVG/Function, AVG/Role, Actual Score/Interviewee

Trust in Master Data

Apart from the MDM director, the scores from all the other participants indicate that currently there is not any trust in the MDM process. This indicates that either the process is not clear to the rest of the business by the MDM Director or that the MDM Director, (even though he tries to put a process in place), has not successfully convinced the users to follow this process.

The following graphs represent the results of the participants. Due to the variety of the participants, different roles, different points of view within the same roles and also due to the small number of samples, each part has been grouped differently. All the answers can be seen in a table format in Appendix 1.

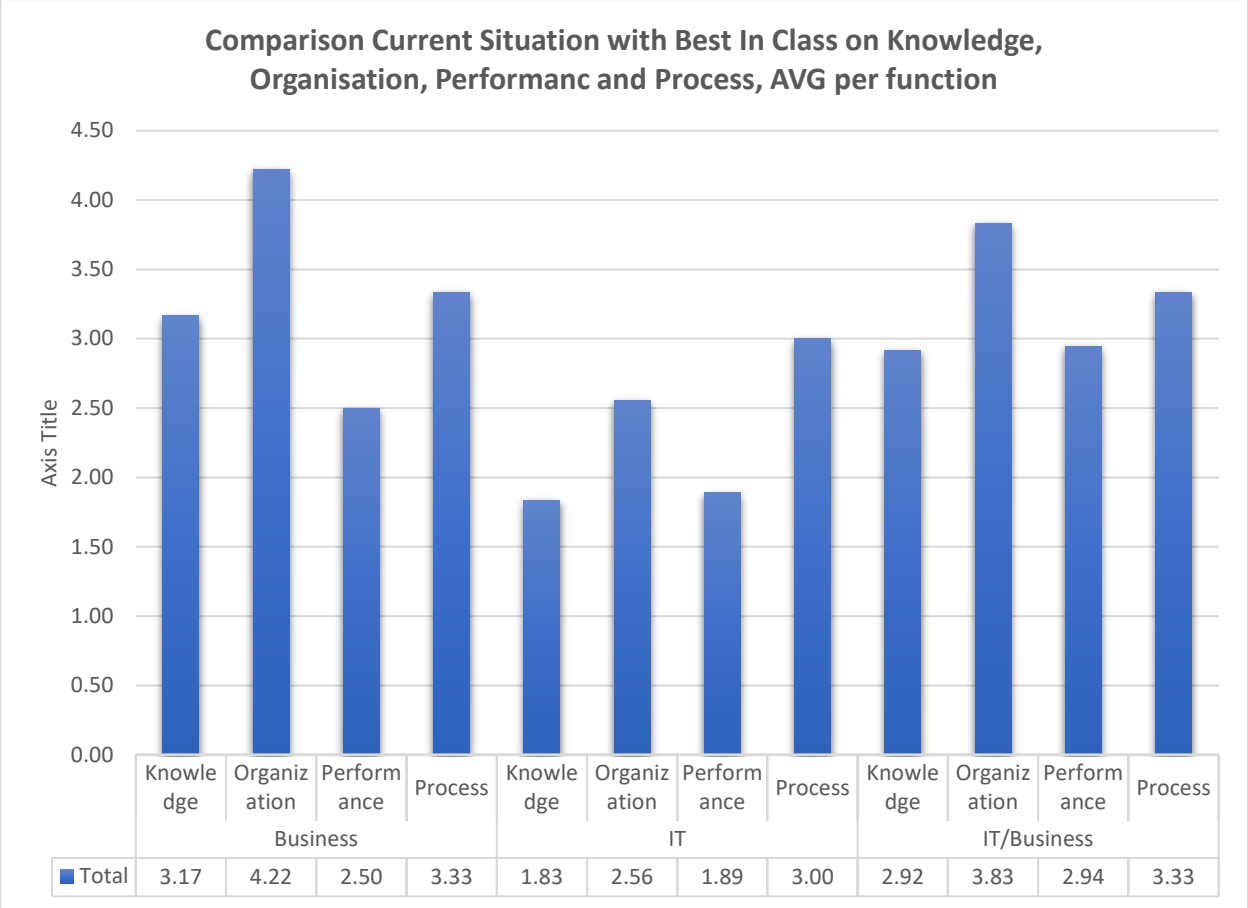


Figure 15 Question 25 - Comparison Current Situation with Best In Class on Knowledge, Organisation, Performance and Process, AVG per function

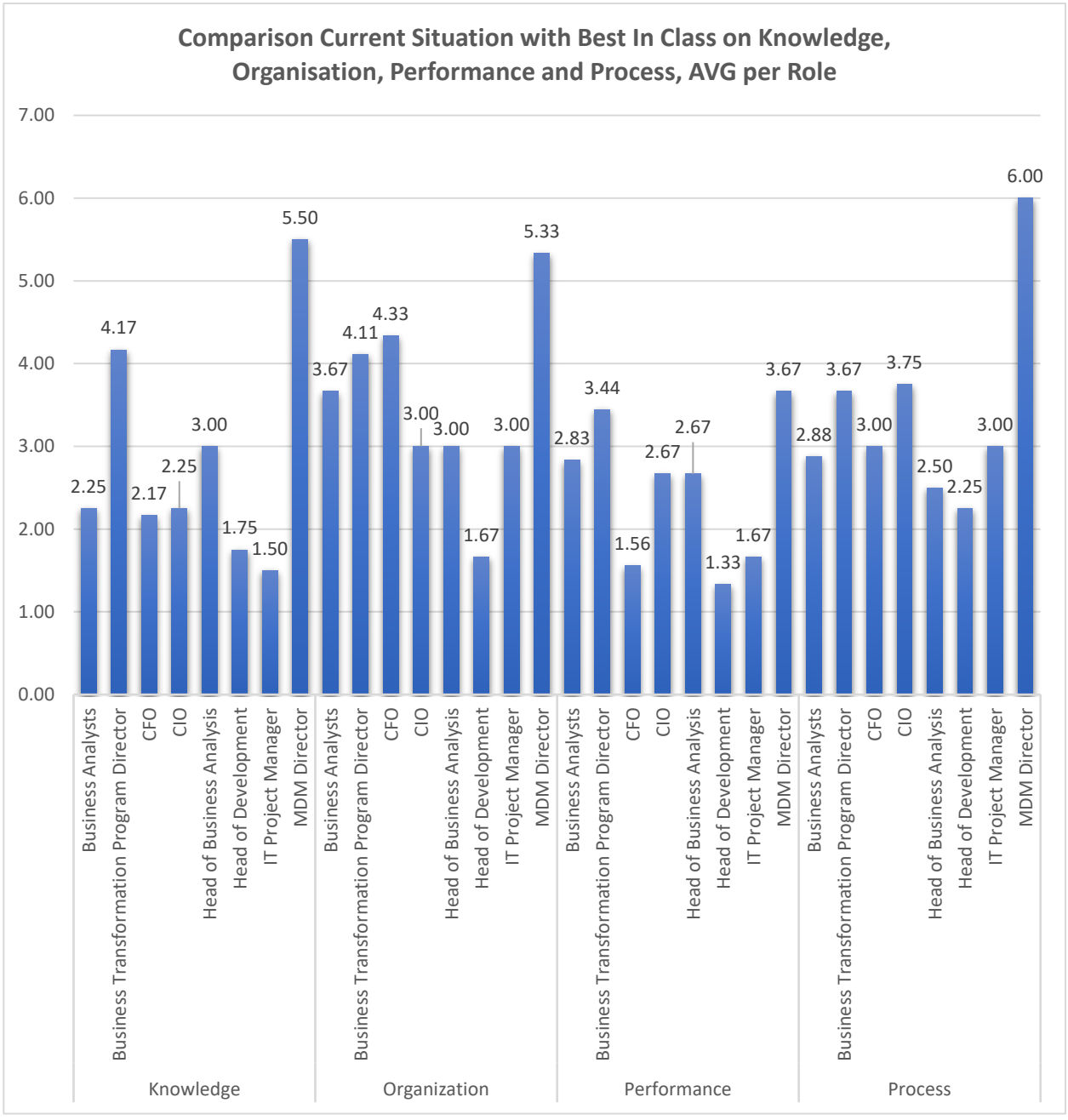


Figure 16 Question 25 - Comparison Current Situation with Best in Class on Knowledge, Organisation, Performance and Process, AVG per function

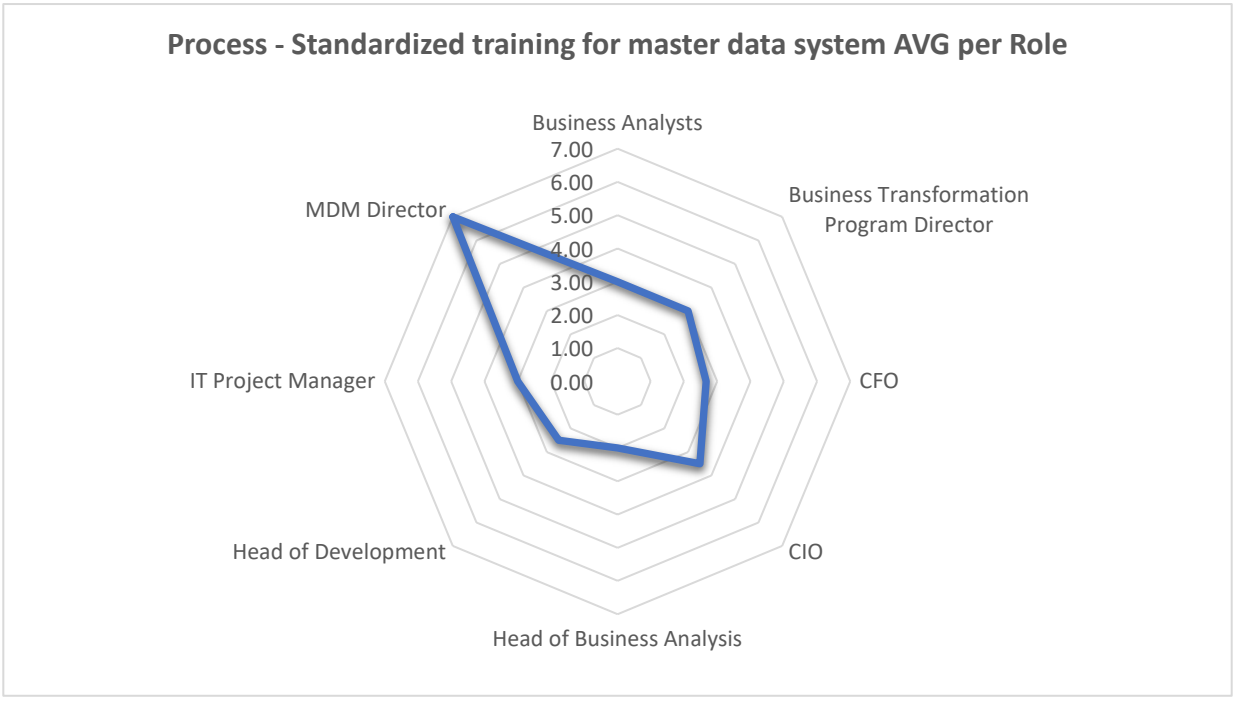


Figure 17 Question 25 - Comparison Current Situation with Best In Class on Process - Standardized training for master data system AVG per Role

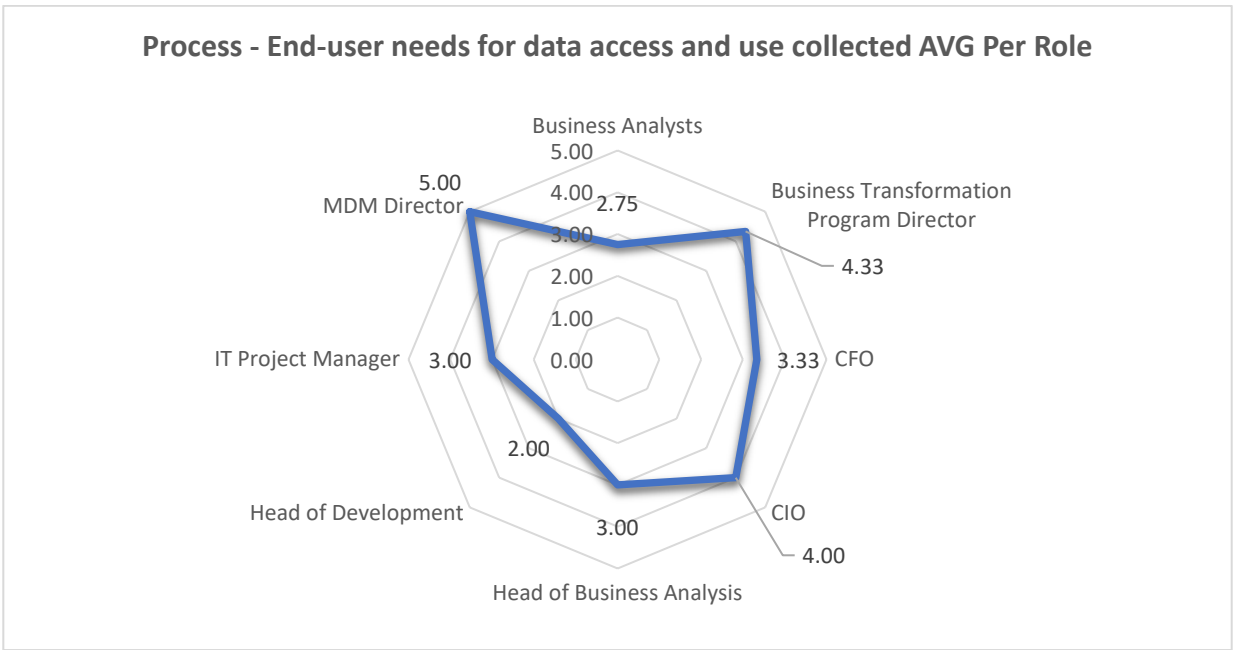


Figure 18 Question 25 - Comparison Current Situation with Best In Class on Process - End-user needs for data access and use collected AVG Per Role

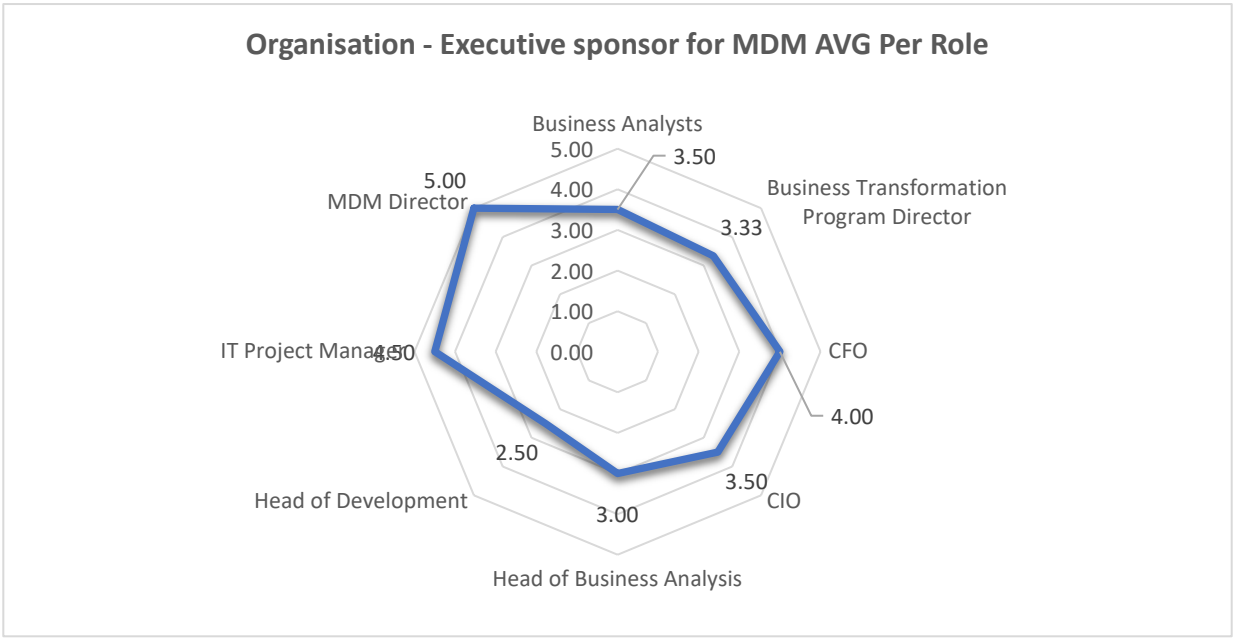


Figure 19 Question 25 - Comparison Current Situation with Best In Class on Organisation - Executive sponsor for MDM AVG Per Role

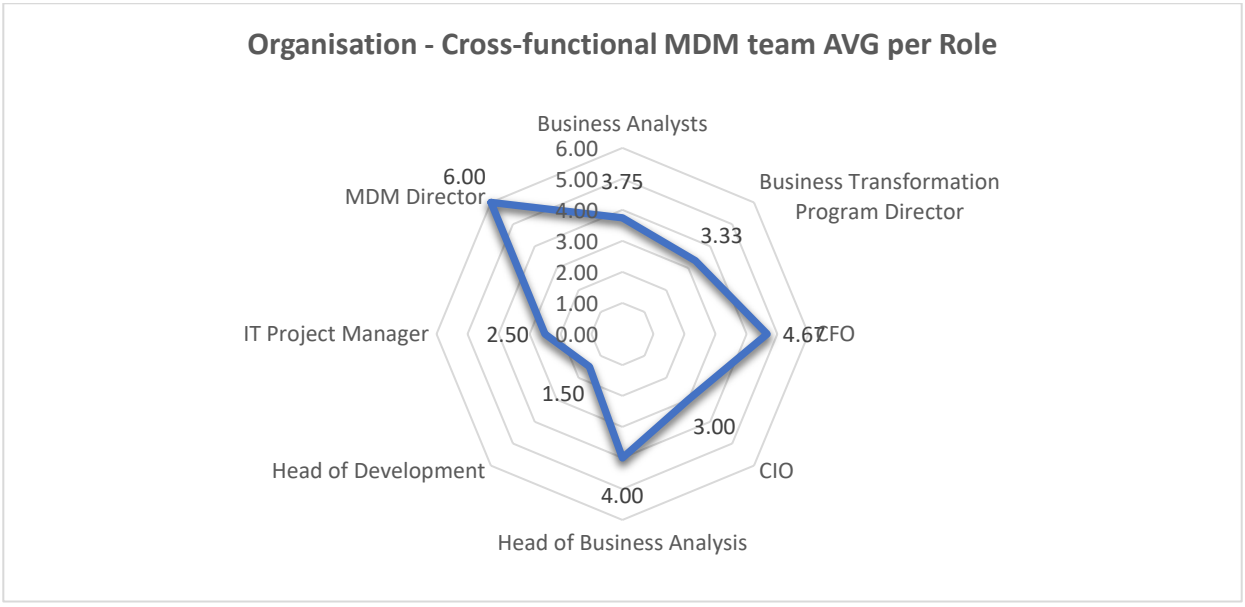


Figure 20 Question 25 - Comparison Current Situation with Best In Class on Organisation - Cross-functional MDM team AVG per Role

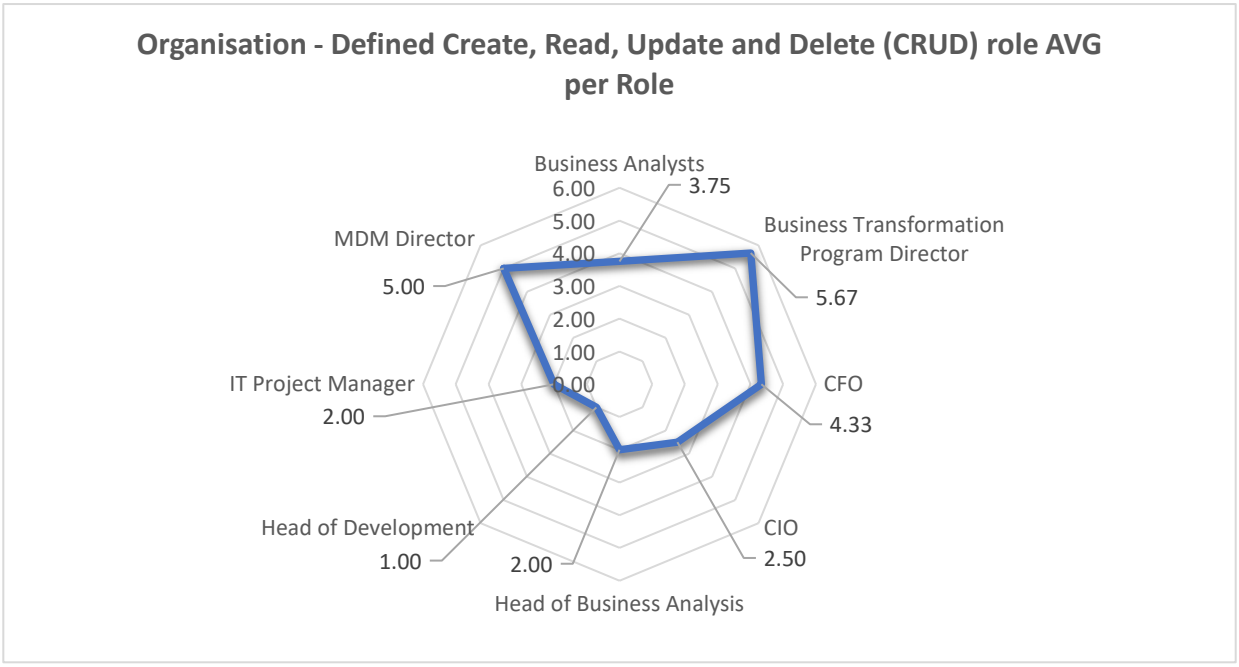


Figure 21. *Figure 21 Question 25 - Comparison Current Situation with Best In Class on Organisation - Defined Create, Read, Update and Delete (CRUD) role AVG per Role*

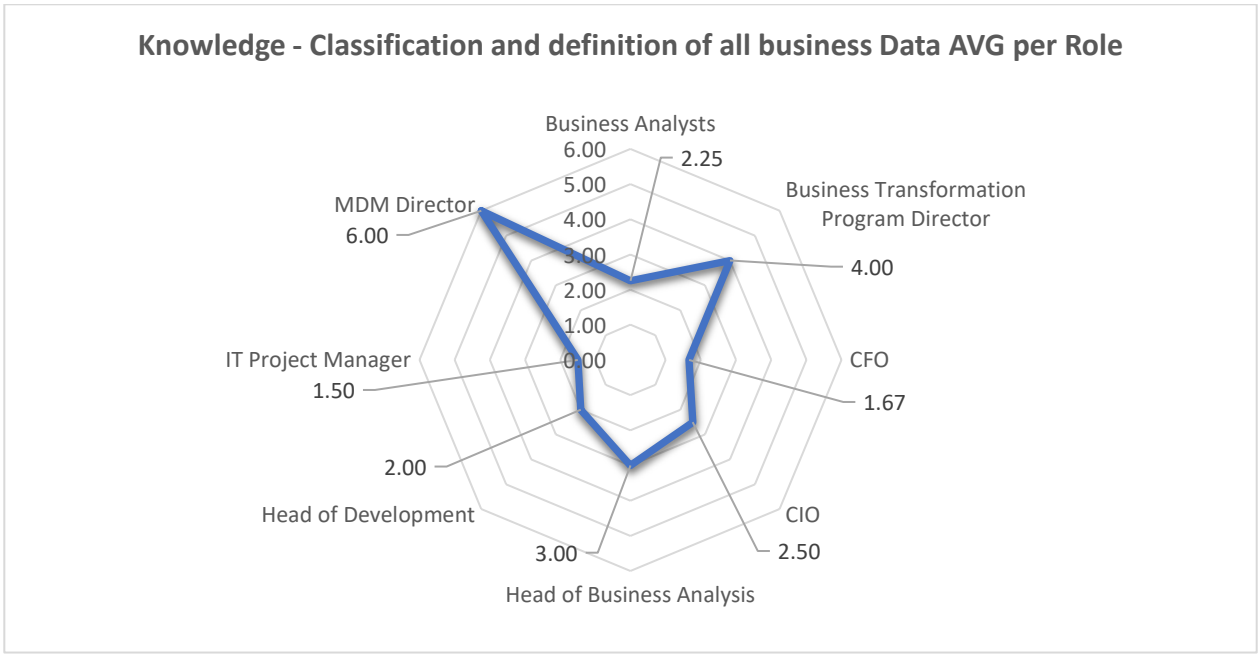


Figure 22 Question 25 - Comparison Current Situation with Best In Class on Organisation - Defined Create, Read, Update and Delete (CRUD) role AVG per Role

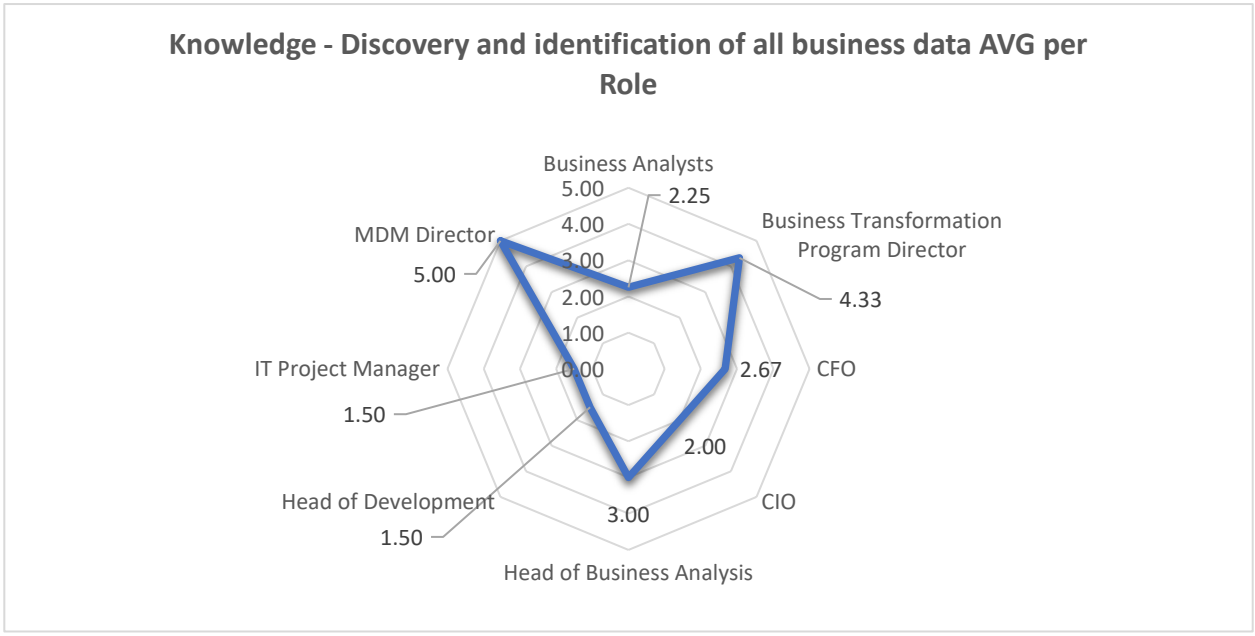


Figure 23 Question 25 - Comparison Current Situation with Best In Class on Knowledge - Discovery and identification of all business data AVG per Role

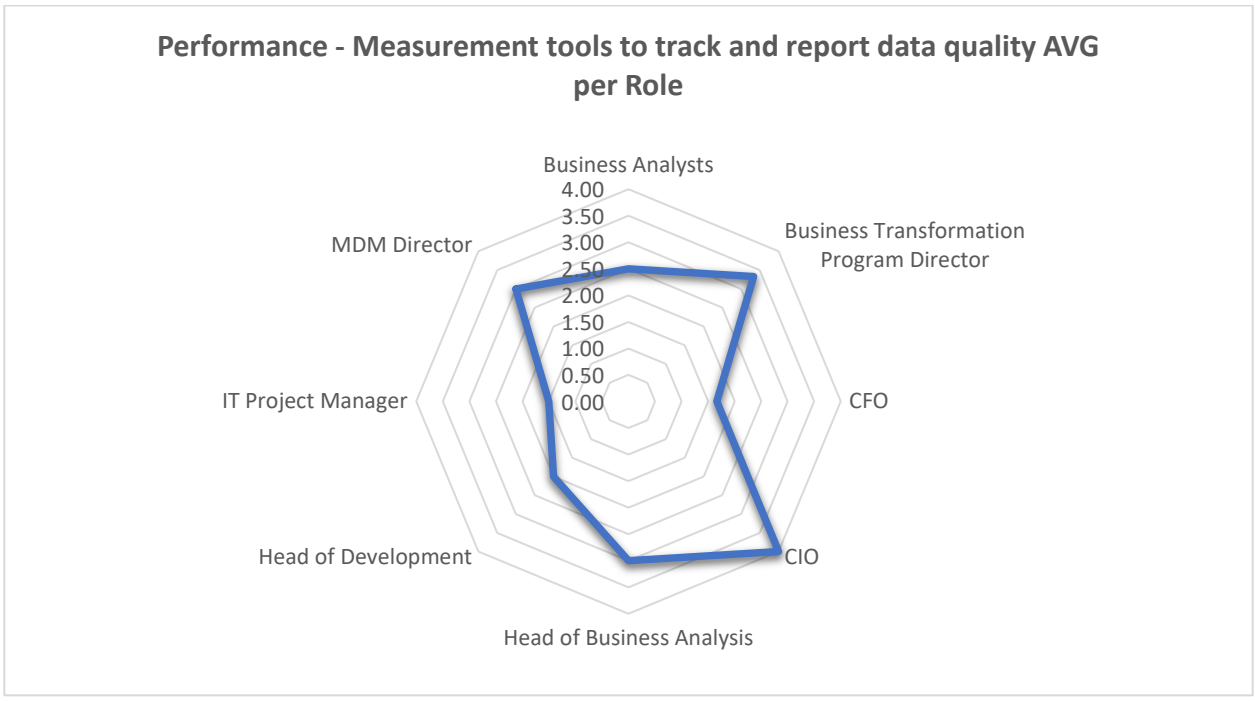


Figure 24 Question 25 - Comparison Current Situation with Best In Class on Performance - Measurement tools to track and report data quality AVG per Role

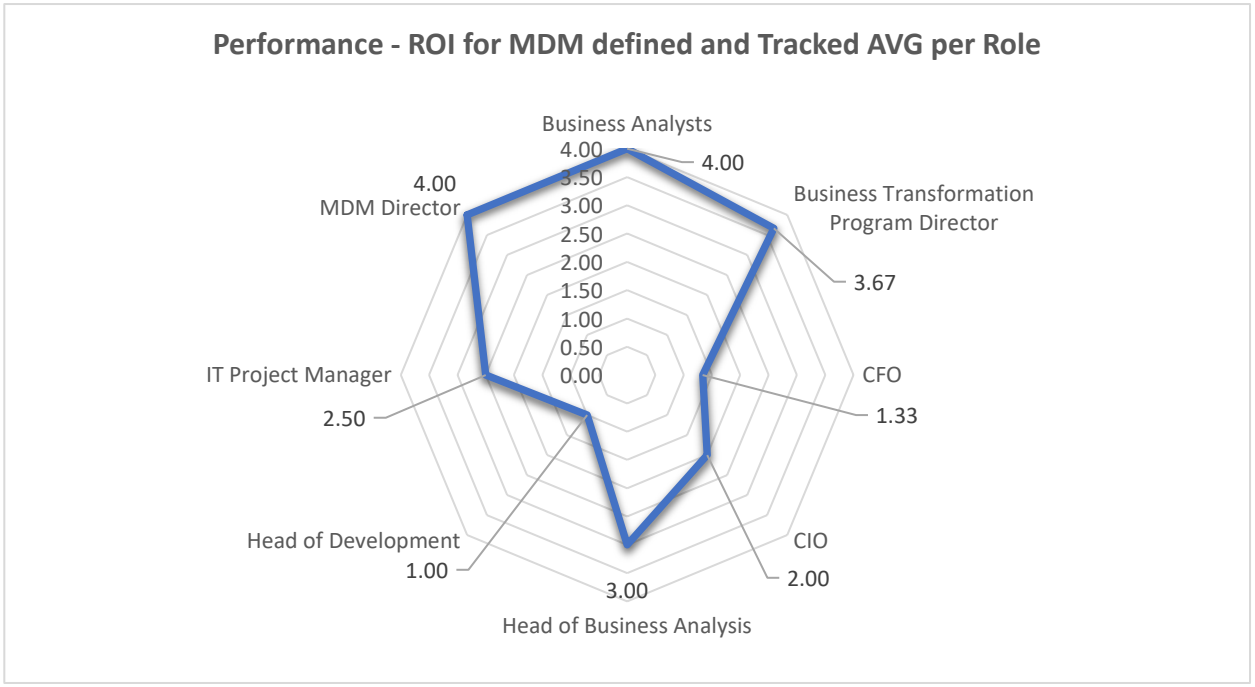


Figure 25 Question 25 - Comparison Current Situation with Best In Class on Performance - ROI for MDM defined and Tracked AVG per Role

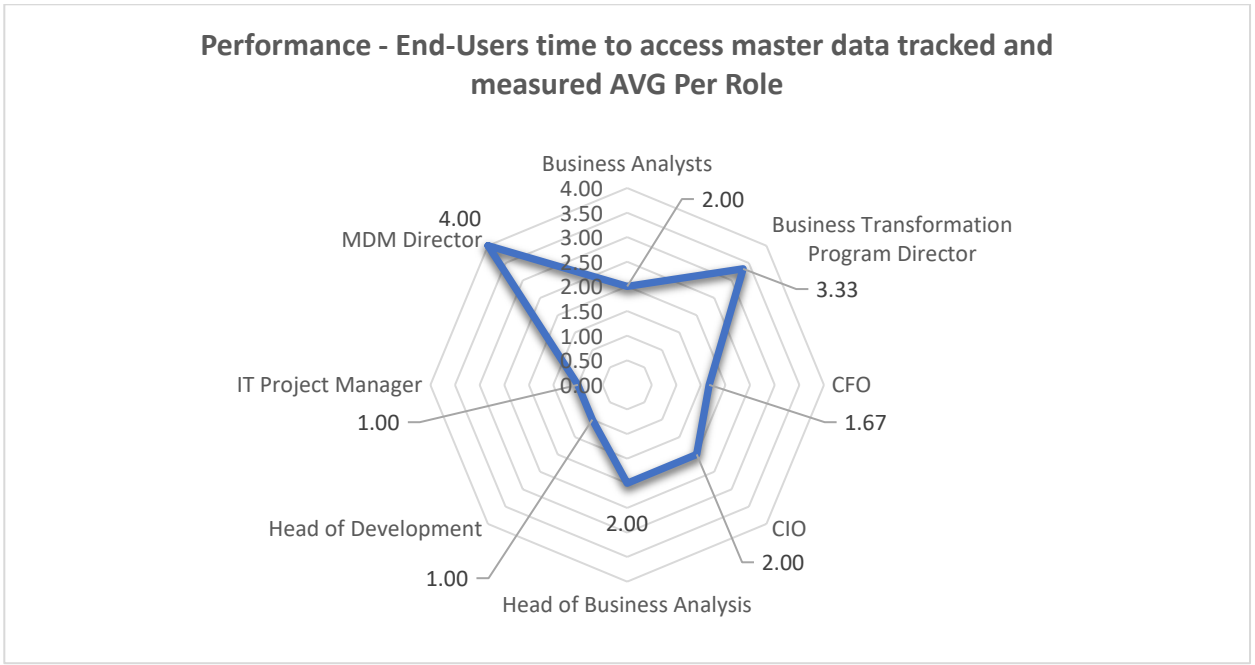


Figure 26 Question 25 - Comparison Current Situation with Best In Class on Performance - End-Users time to access master data tracked and measured AVG Per Role

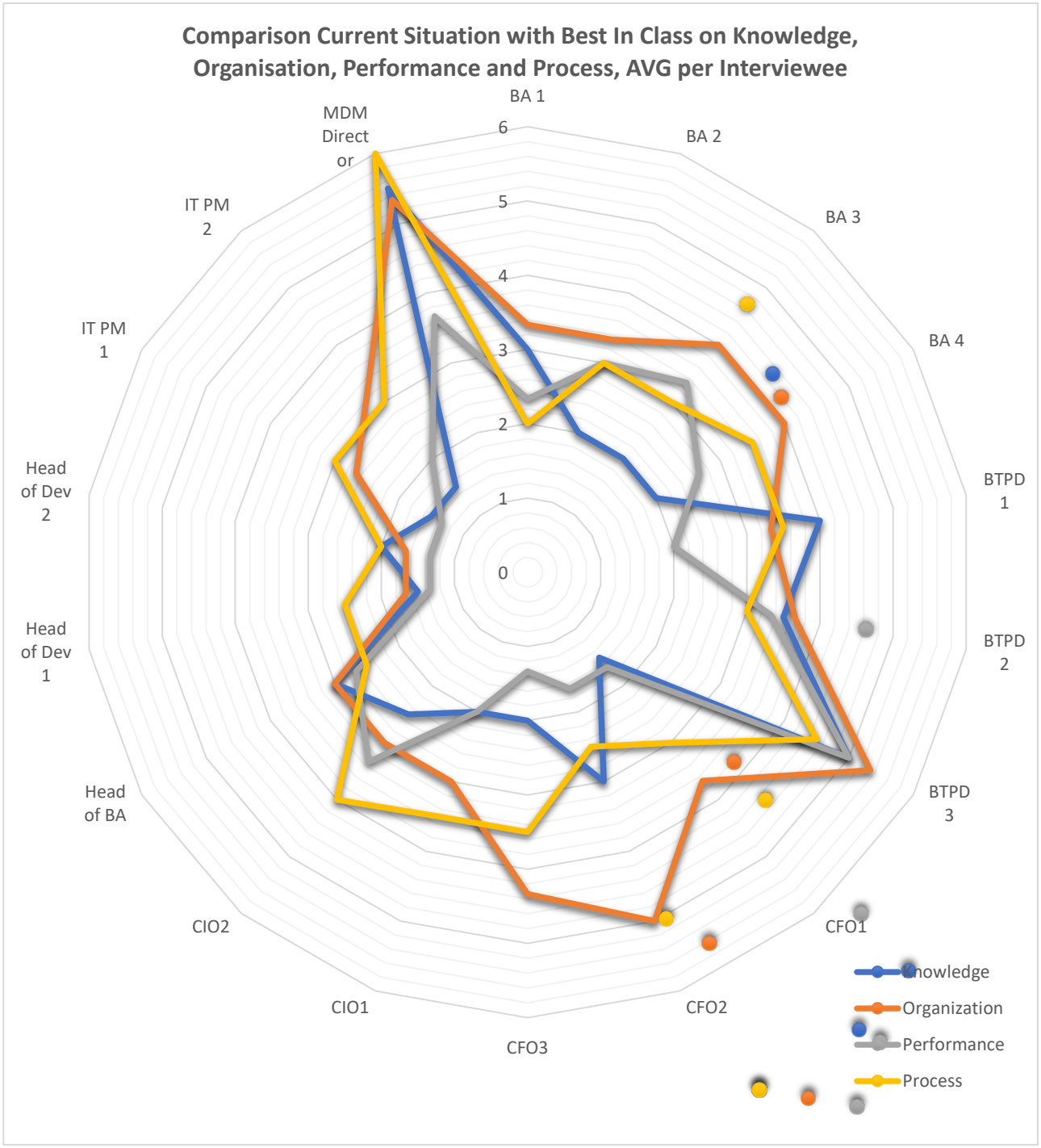
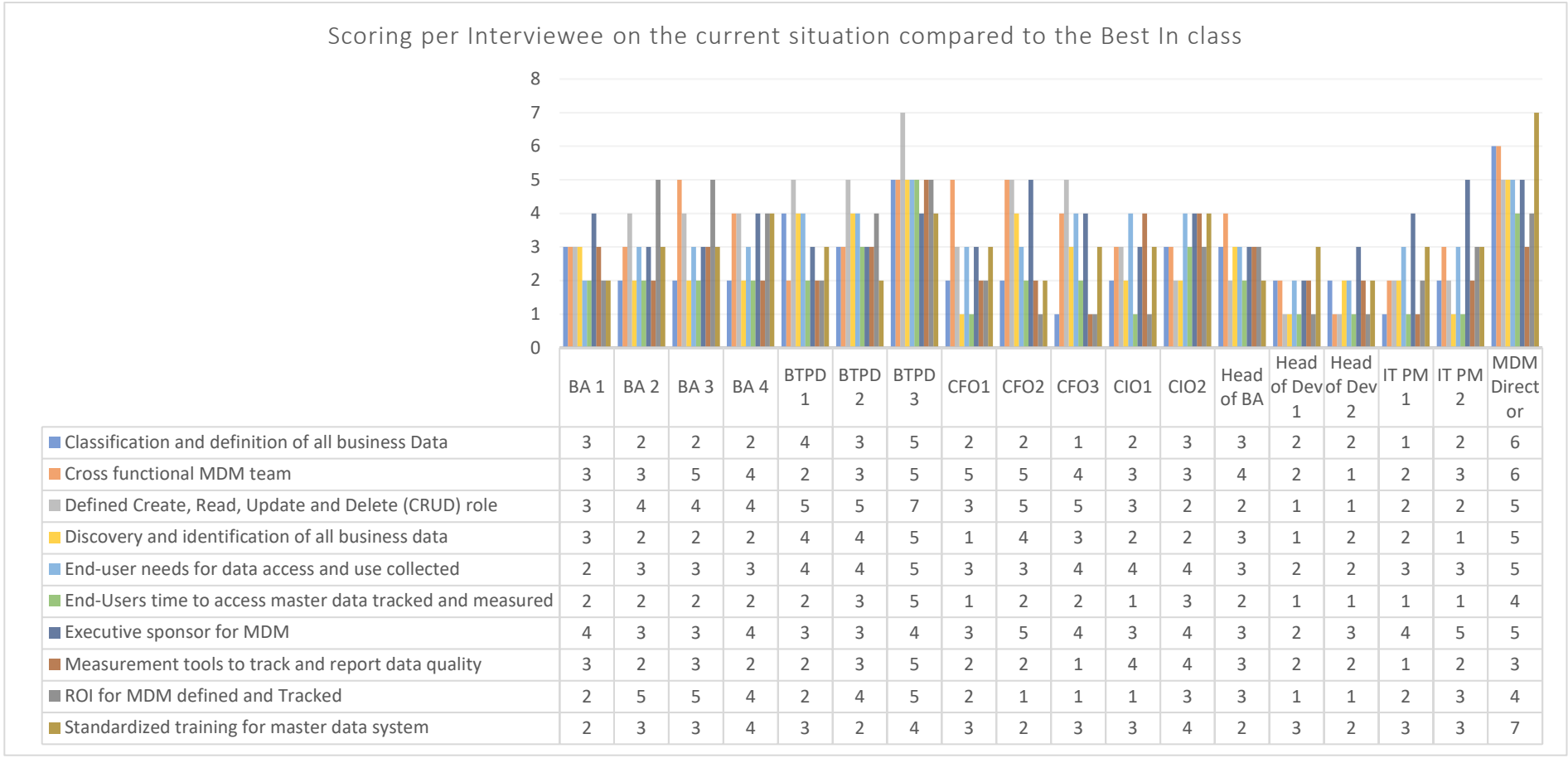


Figure 27 Question 25 - Comparison Current Situation with Best In Class on Knowledge, Organisation, Performance and Process, AVG per Interviewee



	BA 1	BA 2	BA 3	BA 4	BTPD 1	BTPD 2	BTPD 3	CFO1	CFO2	CFO3	CIO1	CIO2	Head of BA	Head of Dev 1	Head of Dev 2	IT PM 1	IT PM 2	MDM Director
Classification and definition of all business Data	3	2	2	2	4	3	5	2	2	1	2	3	3	2	2	1	2	6
Cross functional MDM team	3	3	5	4	2	3	5	5	5	4	3	3	4	2	1	2	3	6
Defined Create, Read, Update and Delete (CRUD) role	3	4	4	4	5	5	7	3	5	5	3	2	2	1	1	2	2	5
Discovery and identification of all business data	3	2	2	2	4	4	5	1	4	3	2	2	3	1	2	2	1	5
End-user needs for data access and use collected	2	3	3	3	4	4	5	3	3	4	4	4	3	2	2	3	3	5
End-Users time to access master data tracked and measured	2	2	2	2	2	3	5	1	2	2	1	3	2	1	1	1	1	4
Executive sponsor for MDM	4	3	3	4	3	3	4	3	5	4	3	4	3	2	3	4	5	5
Measurement tools to track and report data quality	3	2	3	2	2	3	5	2	2	1	4	4	3	2	2	1	2	3
ROI for MDM defined and Tracked	2	5	5	4	2	4	5	2	1	1	1	3	3	1	1	2	3	4
Standardized training for master data system	2	3	3	4	3	2	4	3	2	3	3	4	2	3	2	3	3	7

Figure 28 Question 25 - Scoring per Interviewee on the current situation compared to the Best In-class

Trust in Current and future Process

Based on the answers, there is no trust in the current process. The positive fact, however, is that nearly everybody believes that the future environment will improve the business. None of the interviewees believes in any mobile usage for the MDM process. The optimism that the participants have is not guaranteed and there is a possibility that the participants will face challenges after the implementation. The reason is that most of the interviewees believe that an MDM solution, should automatically resolve all the data issues. However, nearly all the participants are focused on the **Registry** (Obenhofer and Dreibelbis 2008) (Chapter 3.4.2.) implementation style and some of them to **Coexistence** (Obenhofer and Dreibelbis 2008) (Chapter 3.4.2.) implementation style having to face challenges with the systems and what kind of services can be implemented. This level of optimism does not comply with what is feasible in real terms. Based on the conversations, the expectation and the belief are targeting a **Transactional** (Obenhofer and Dreibelbis 2008) (Chapter 3.4.2.) style implementation however, this is not feasible in many cases.

The following graphs represent the results of the participants. Due to the variety of the participants, different roles, different points of view within the same roles and due to the small number of samples, each part has been grouped differently. All the answers can be seen in a table format in Appendix 1.

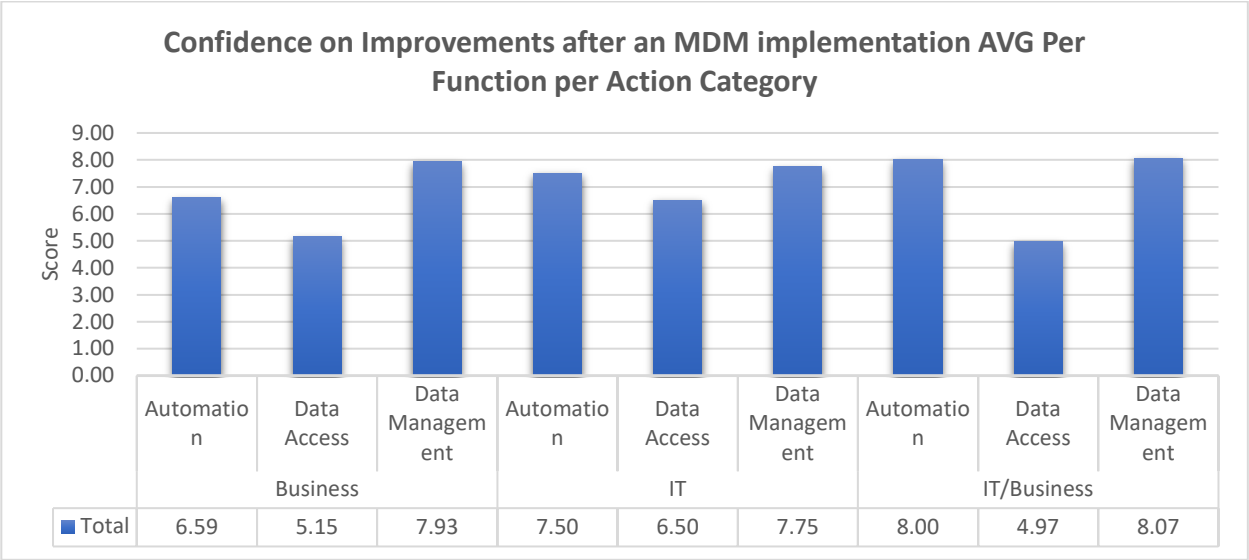


Figure 29 Question 27 - Confidence on Improvements after an MDM implementation AVG Per Function per Action Category

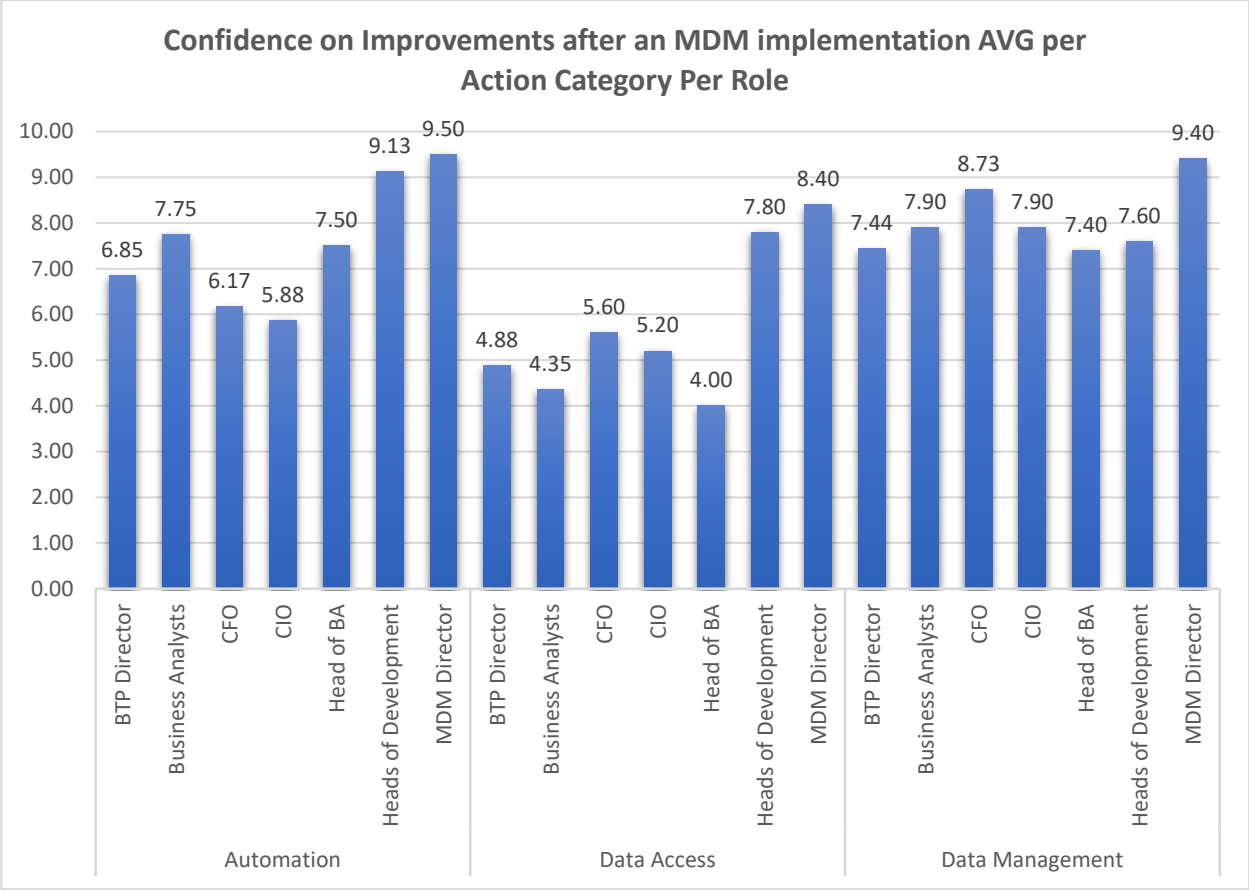


Figure 30 Question 27 - Confidence on Improvements after an MDM implementation AVG per Action Category Per Role

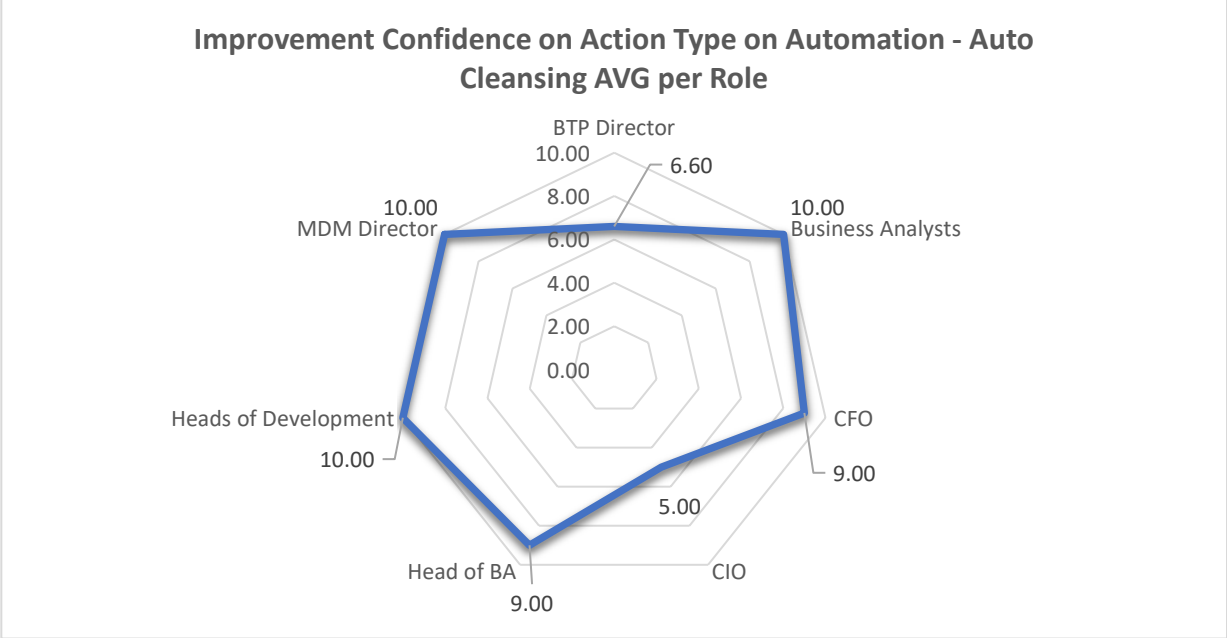


Figure 31 Question 27 - Improvement Confidence on Action Type on Automation - Auto Cleansing AVG per Role

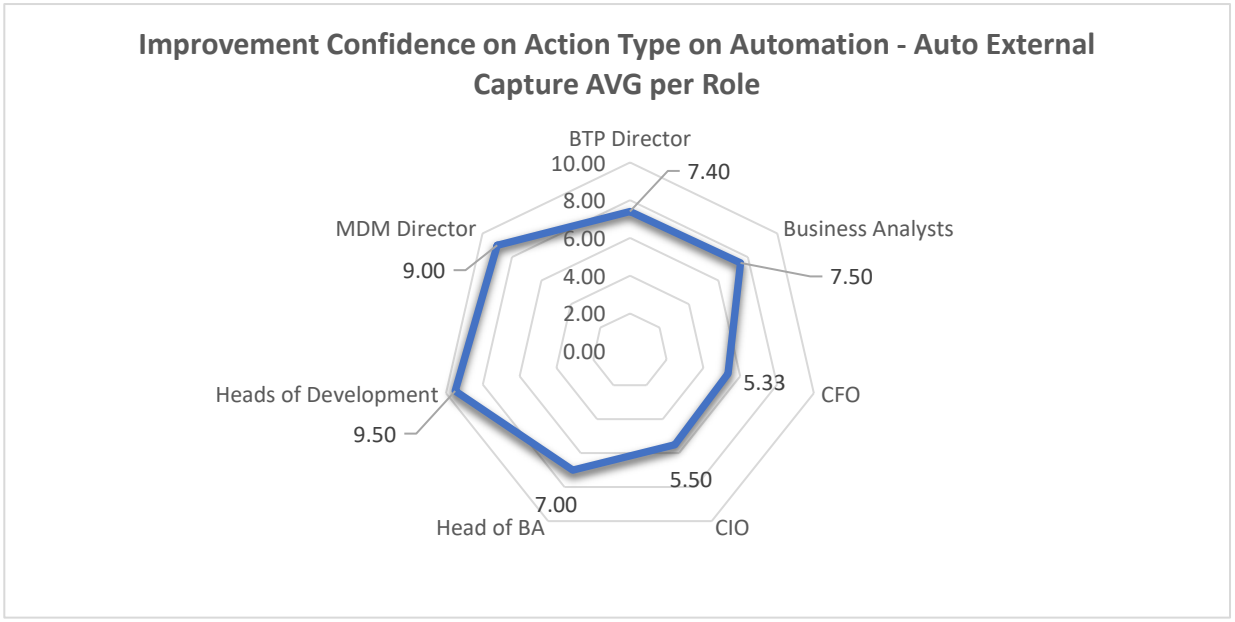


Figure 32 Question 27 - Improvement Confidence on Action Type on Automation - Auto External Capture AVG per Role

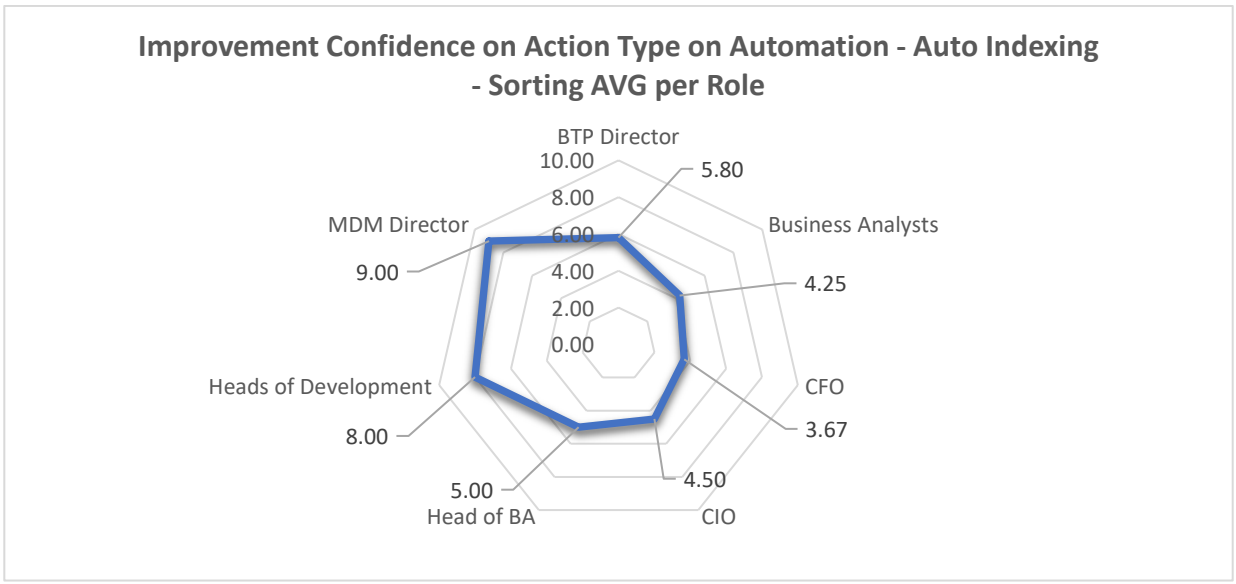


Figure 33 Question 27 - Improvement Confidence on Action Type on Automation - Auto Indexing - Sorting AVG per Role

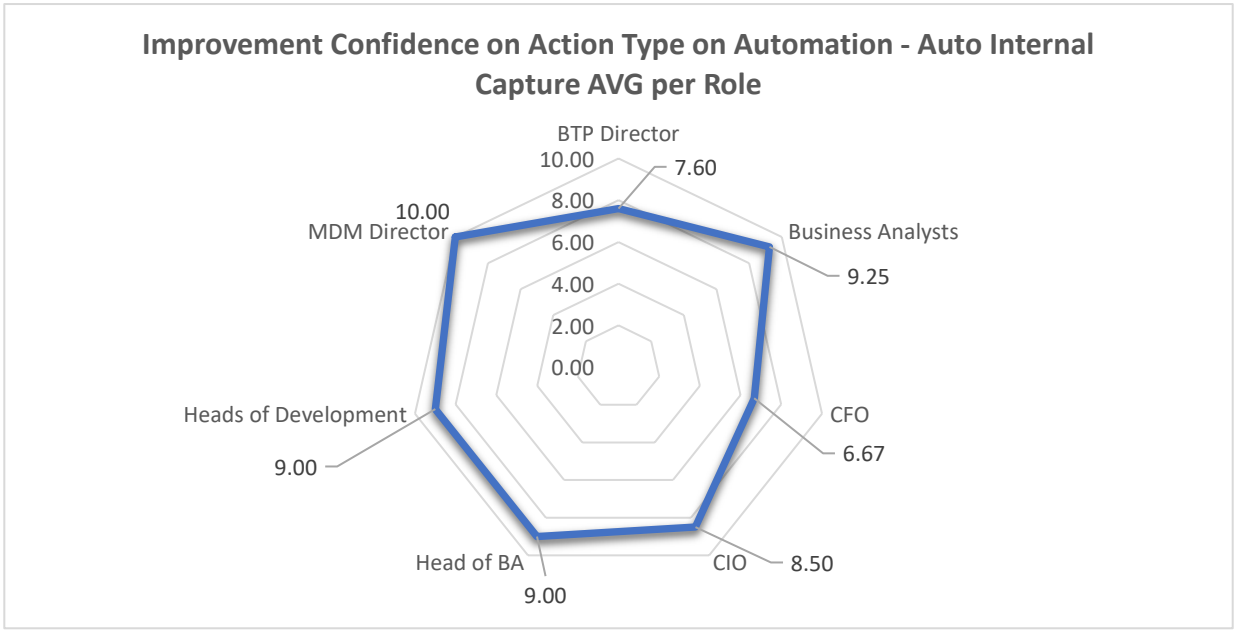


Figure 34 Question 27 - Improvement Confidence on Action Type on Automation - Auto Internal Capture AVG per Role

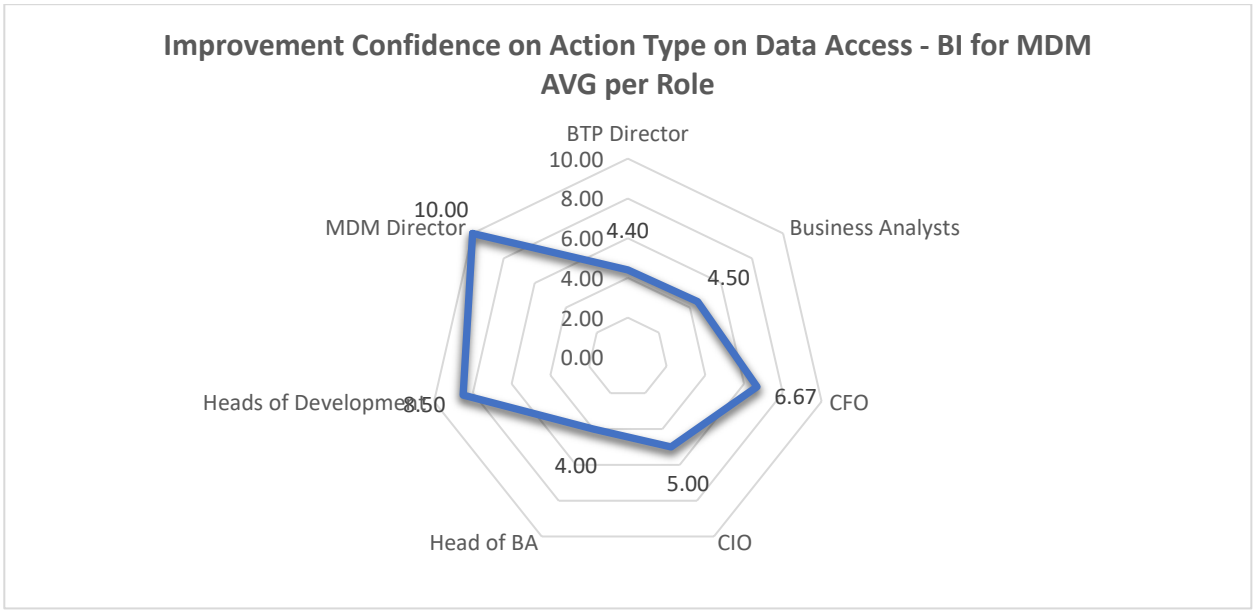


Figure 35 Question 27 - Improvement Confidence on Action Type on Data Access - BI for MDM AVG per Role

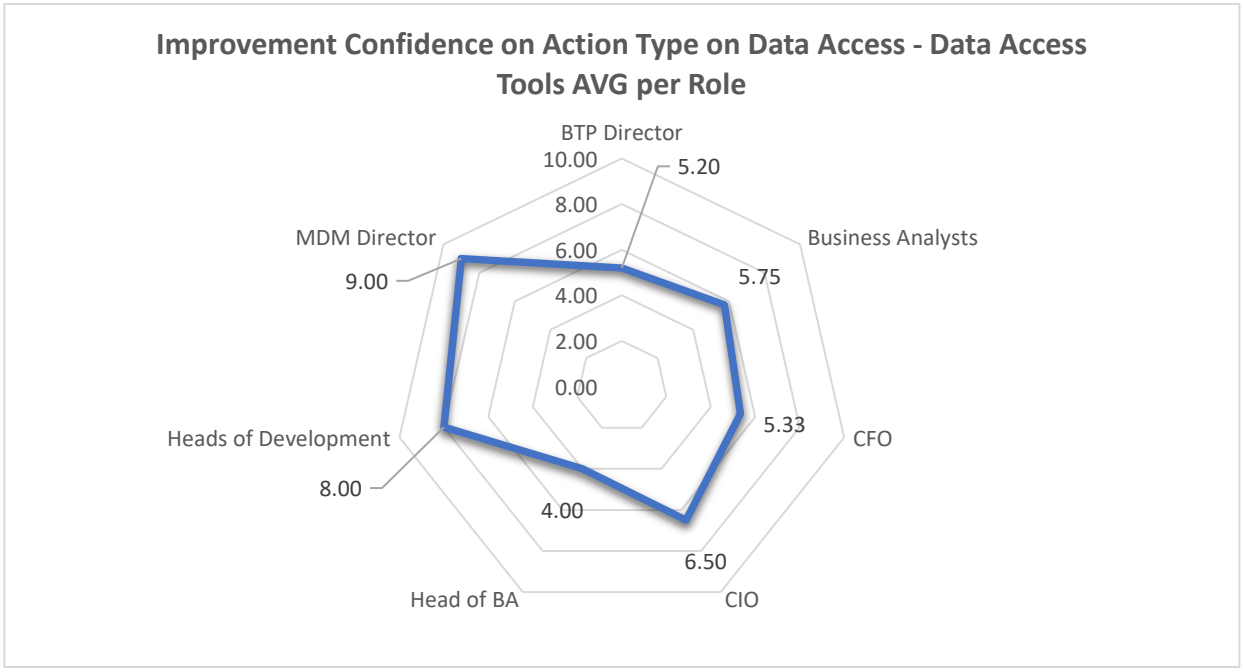


Figure 36 Question 27 - Improvement Confidence on Action Type on Data Access - Data Access Tools AVG per Role

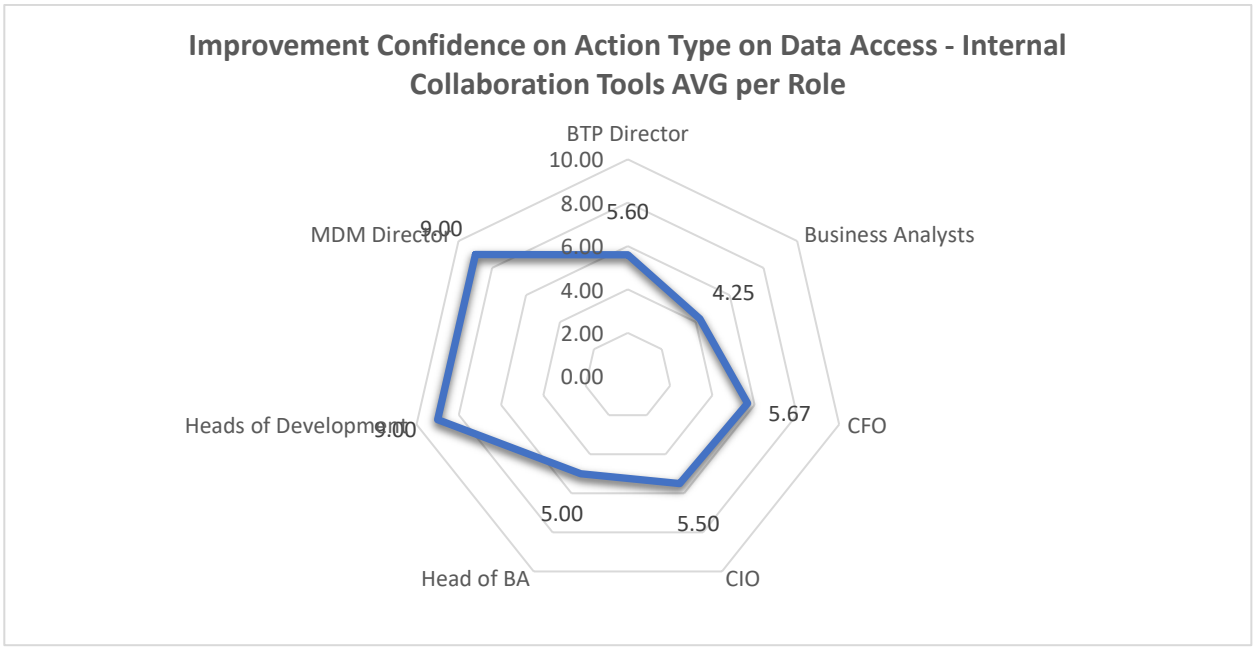


Figure 37 Question 27 - Improvement Confidence on Action Type on Data Access - Data Access Tools AVG per Role

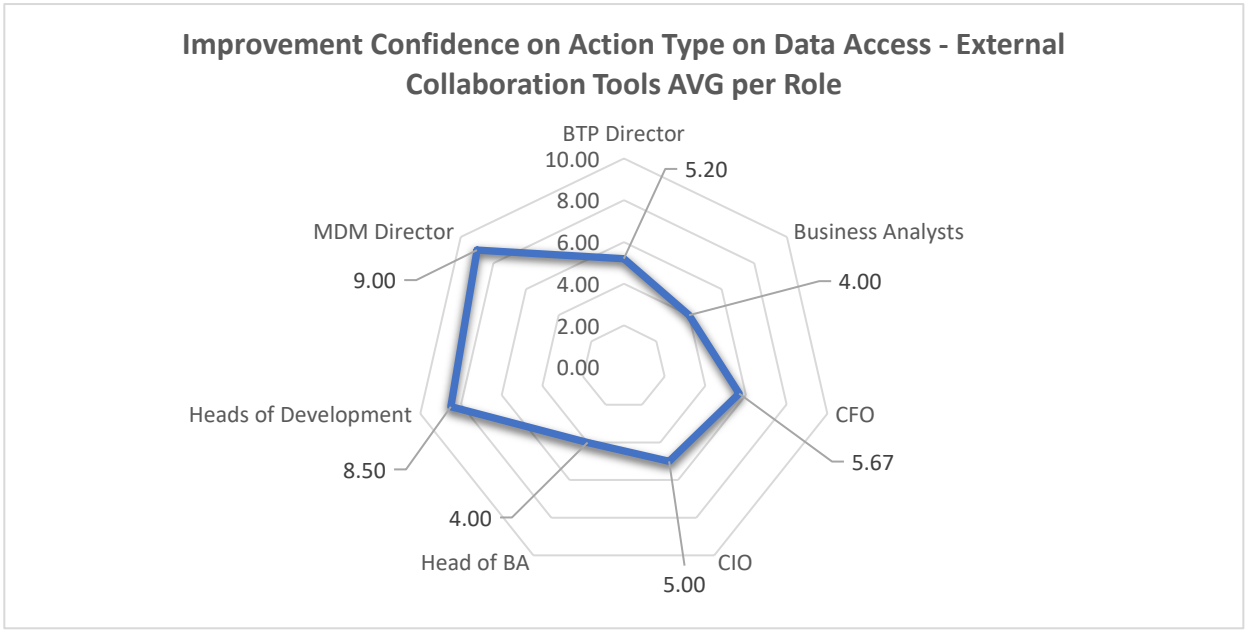


Figure 38 Question 27 - Improvement Confidence on Action Type on Data Access - Data Access Tools AVG per Role

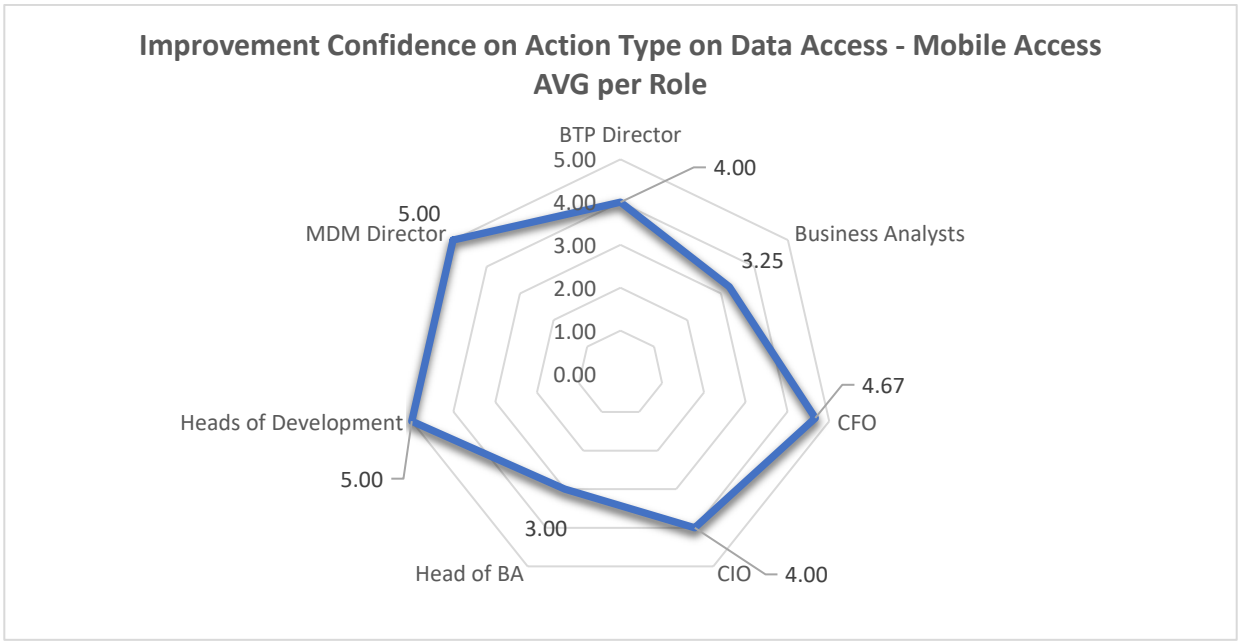


Figure 39 Question 27 - Improvement Confidence on Action Type on Data Access - Mobile Access AVG per Role

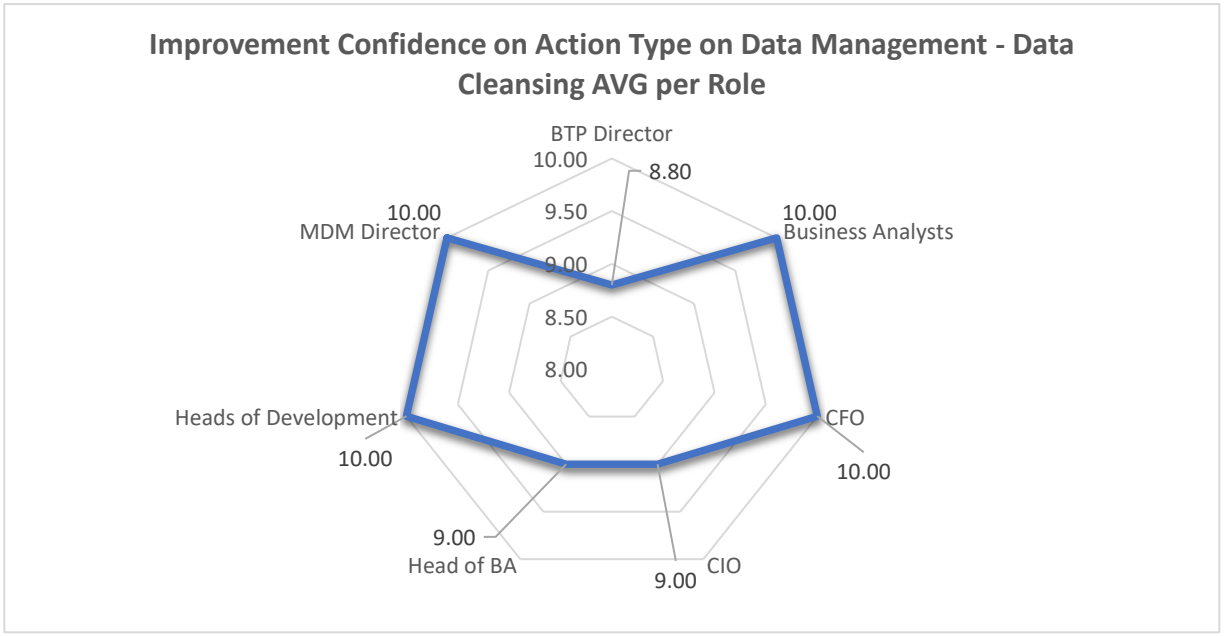


Figure 40 Question 27 - Improvement Confidence on Action Type on Data Management - Data Cleansing AVG per Role

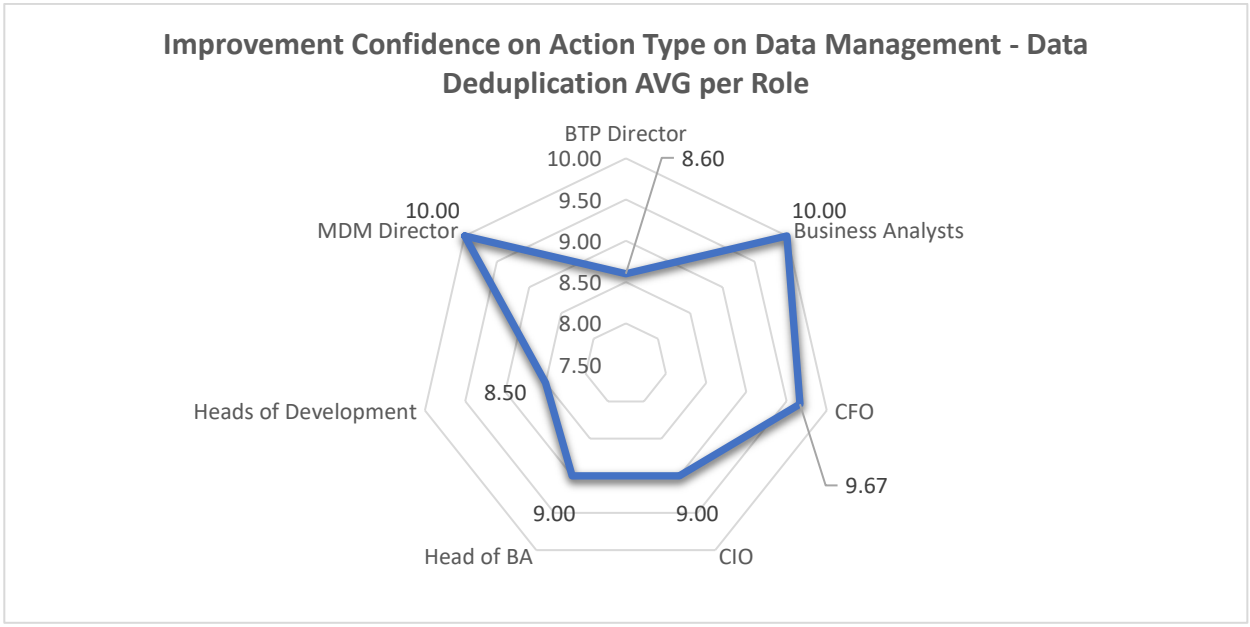


Figure 41 Question 27 - Improvement Confidence on Action Type on Data Management - Data Deduplication AVG per Role

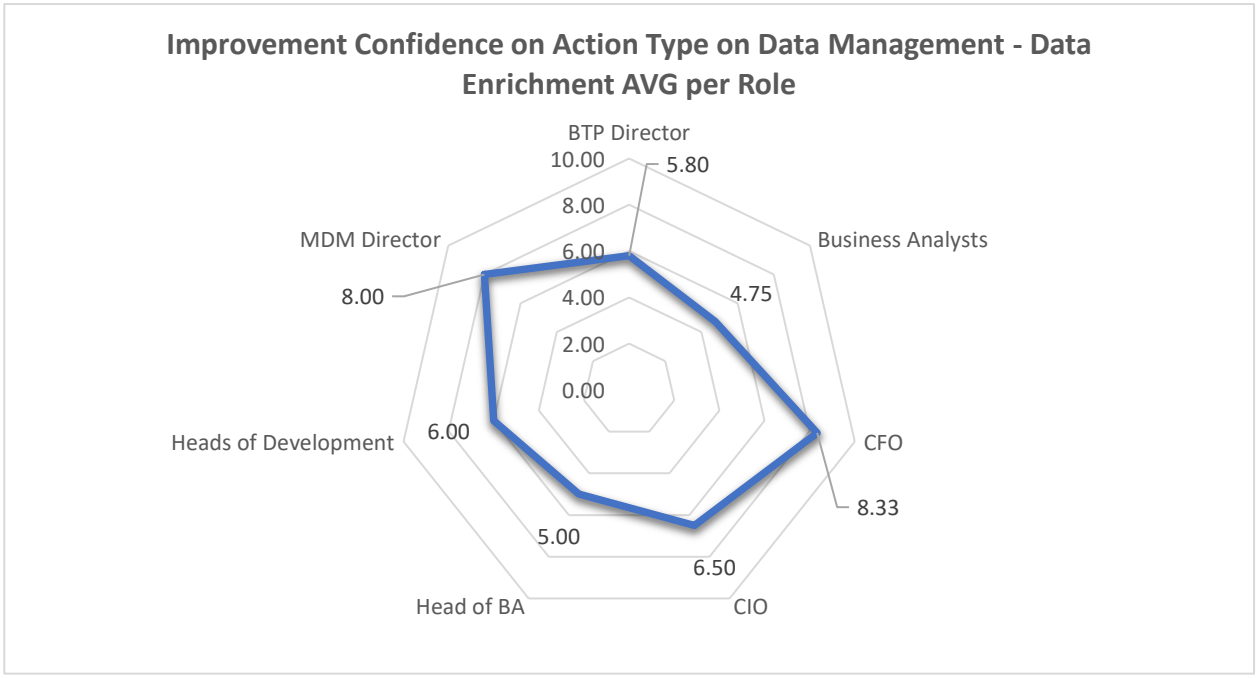


Figure 42 Question 27 - Improvement Confidence on Action Type on Data Management - Data Enrichment AVG per Role

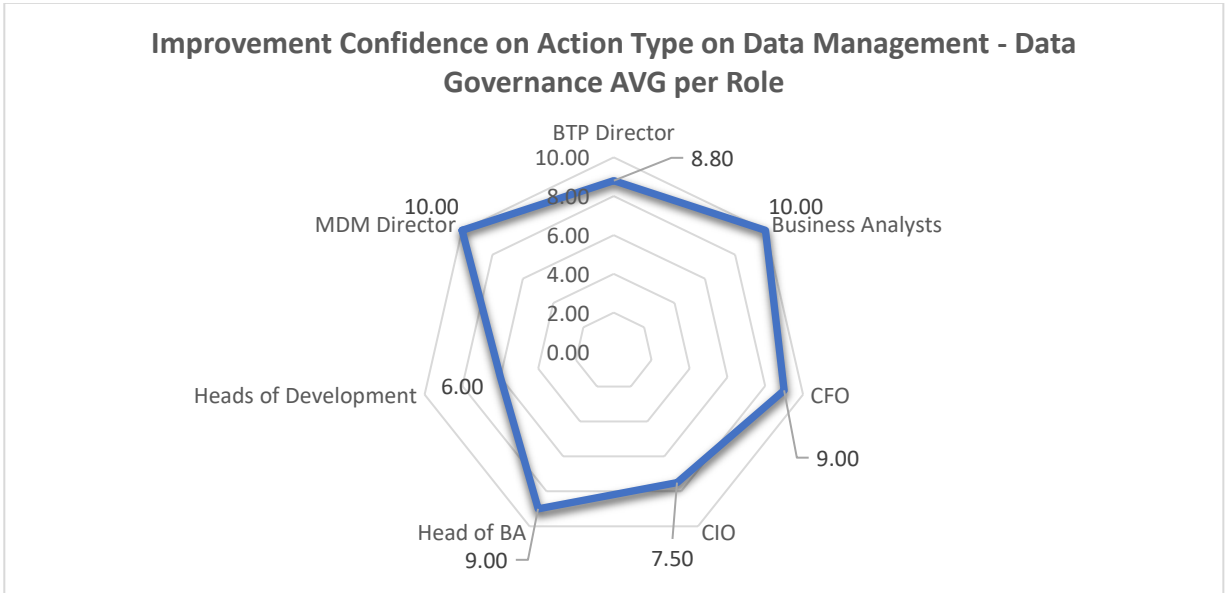


Figure 43 Question 27 - Improvement Confidence on Action Type on Data Management - Data Governance AVG per Role

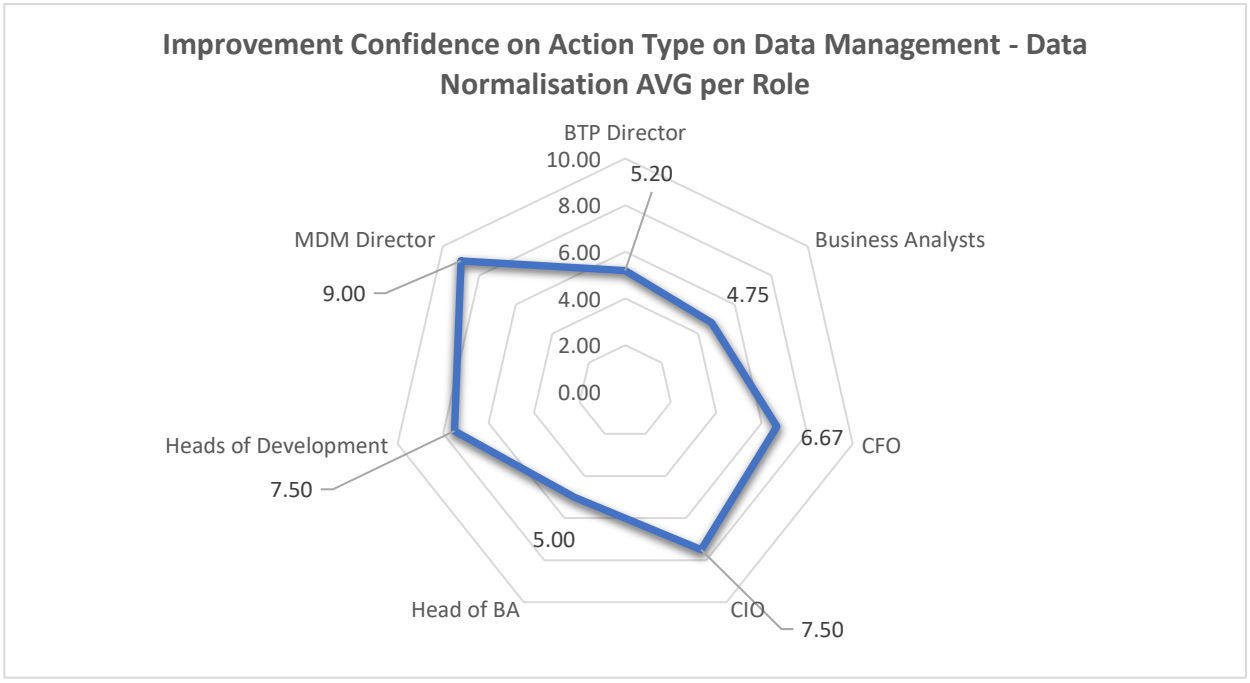


Figure 44 Question 27 - Improvement Confidence on Action Type on Data Management - Data Normalisation AVG per Role

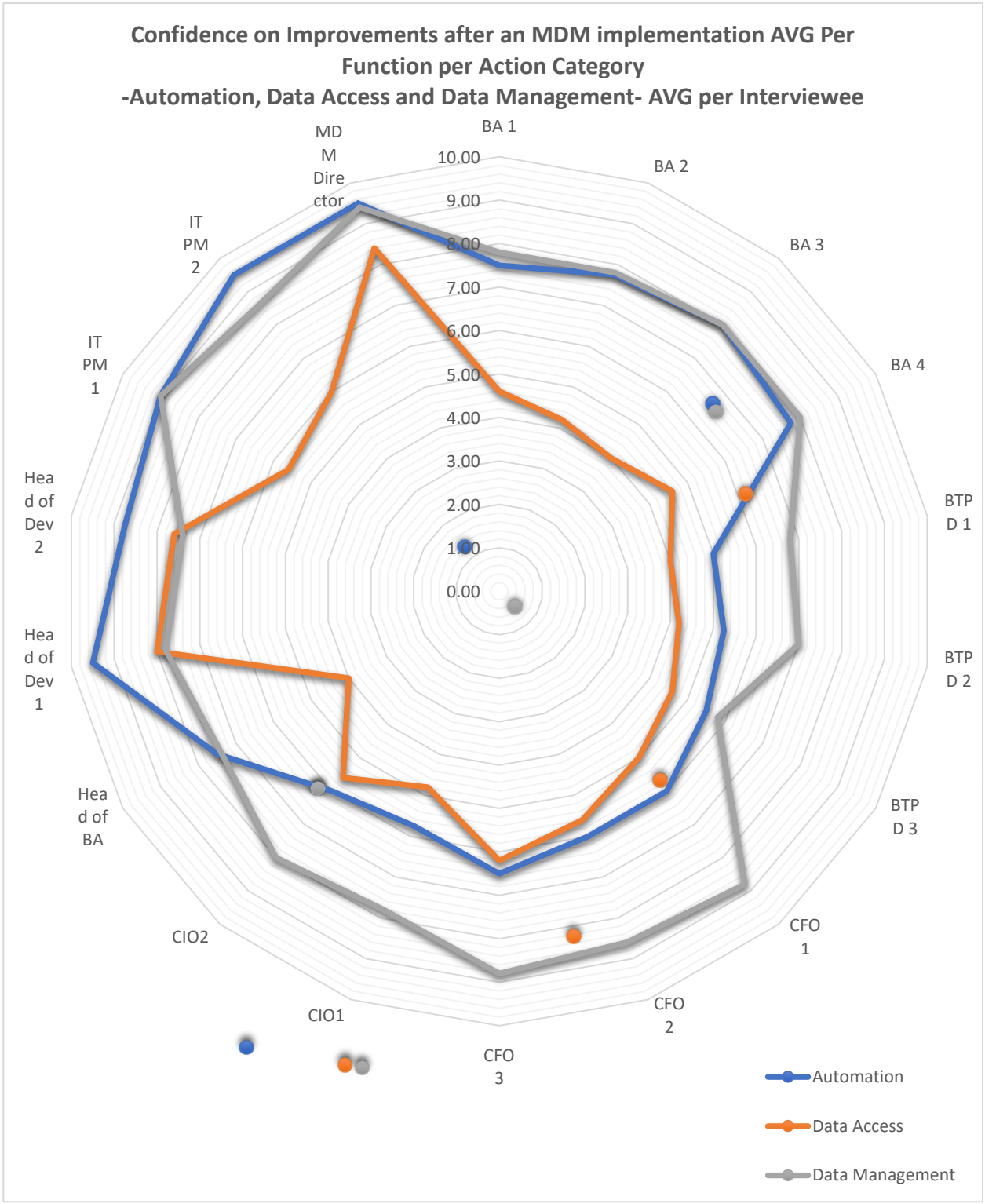
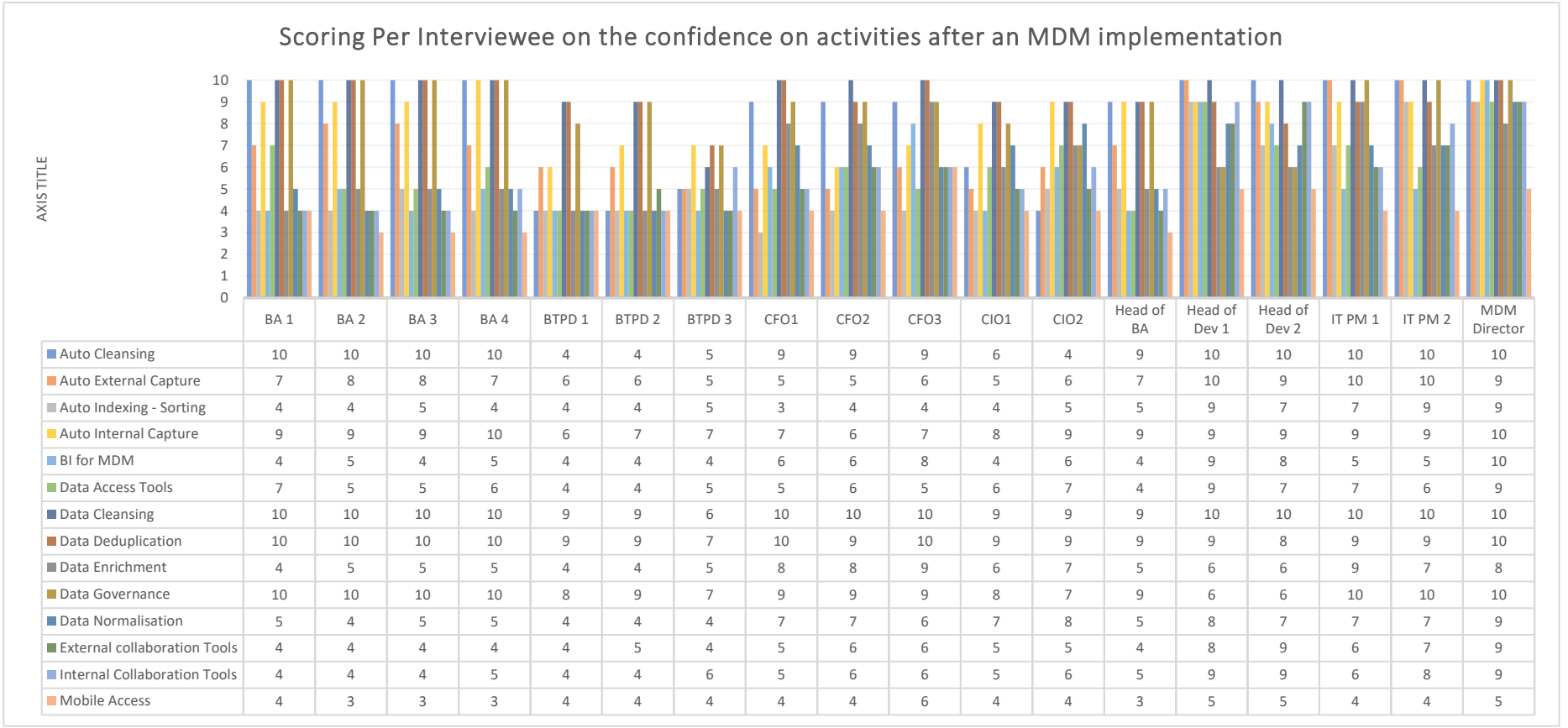


Figure 45 Question 27 - Confidence on Improvements after an MDM implementation AVG Per Function per Action Category



	BA 1	BA 2	BA 3	BA 4	BTPD 1	BTPD 2	BTPD 3	CFO1	CFO2	CFO3	CIO1	CIO2	Head of BA	Head of Dev 1	Head of Dev 2	IT PM 1	IT PM 2	MDM Director
Auto Cleansing	10	10	10	10	4	4	5	9	9	9	6	4	9	10	10	10	10	10
Auto External Capture	7	8	8	7	6	6	5	5	5	6	5	6	7	10	9	10	10	9
Auto Indexing - Sorting	4	4	5	4	4	4	5	3	4	4	4	5	5	9	7	7	9	9
Auto Internal Capture	9	9	9	10	6	7	7	7	6	7	8	9	9	9	9	9	9	10
BI for MDM	4	5	4	5	4	4	4	6	6	8	4	6	4	9	8	5	5	10
Data Access Tools	7	5	5	6	4	4	5	5	6	5	6	7	4	9	7	7	6	9
Data Cleansing	10	10	10	10	9	9	6	10	10	10	9	9	9	10	10	10	10	10
Data Deduplication	10	10	10	10	9	9	7	10	9	10	9	9	9	9	8	9	9	10
Data Enrichment	4	5	5	5	4	4	5	8	8	9	6	7	5	6	6	9	7	8
Data Governance	10	10	10	10	8	9	7	9	9	9	8	7	9	6	6	10	10	10
Data Normalisation	5	4	5	5	4	4	4	7	7	6	7	8	5	8	7	7	7	9
External collaboration Tools	4	4	4	4	4	5	4	5	6	6	5	5	4	8	9	6	7	9
Internal Collaboration Tools	4	4	4	5	4	4	6	5	6	6	5	6	5	9	9	6	8	9
Mobile Access	4	3	3	3	4	4	4	4	4	6	4	4	3	5	5	4	4	5

Figure 46 Question 27 - Scoring Per Interviewee on the confidence on activities after an MDM implementation

5.4. Audit Reports and Interviews Highlights

After the presentation of the audit reports and the interviews. This section consists of a consolidated view of the results and a list of highlights of the findings. The top 5 highlights are as follows:

1. All the organisations and all the participants are aware that there are multiple **data quality challenges** within their business processes. That is why there is a level of investment to address these challenges and find a resolution. Data quality is important for every function within the business and the longest period a challenge is not resolved, the highest the cost for the organisation overall would be. As it was presented from the interviews and the audits, the cost is not only measured on the number of fines or penalties a company could get from a vendor or a customer. There are costs for every resource that is allocated internally to resolve these challenges. People should disengage from their normal activities, so they can be allocated to a resolution of a data quality issue. As a result, other activities are not progressing in the originally estimated time. Based on that new budgets should be reconsidered and allocated, and the cost is multiplied across all the functions.
2. There is a common theme, that the communication between functions is complicated. In most cases, **change management** is poorly distributed across related functions. As a result, significant changes in business processes are kept within the department that made the changes. Then, there is an assumption that these changes should be assumed by other members or functions and they should be taken into consideration. Also, some processes are kept hidden from other members or different functions because the process owners are not willing to share them. This usually happens because there is not any knowledge of why this process exists in the first instance. Besides, people are reluctant to changes since any change can lead to redundancy of this person's position or weakening of their part in the process.
3. **Senior management commitment**. It appears that the senior management supports the cause and they need the data quality issues to be resolved. However, when the time comes that there is a requirement for this support to be actioned, there multiple obstacles that

require resolution before any action. The main theme is that senior management believes that they actively support the process but the people who try to implement the process suggest that their support is limited. The trust between the teams is a key element that needs to be earned and within an organisation, this trust should be there. If this is not the case, it means that the appropriate people are not the correct positions. This results in more time spending on convincing the senior management of what needs to be done instead of making progress.

4. **Preferred solution**. The results show that the preferred option is an already developed and proven solution. MDM is not a simple IT development project and requires specific skills just for the technicalities of the process. MDM vendors have dedicated teams and research and development departments to optimise the process and to improve the application. Organisations that wish to apply this process and acquire this software are not software vendors and their focus is to increase their revenue by selling more of their products or their services. Having people in leading positions suggesting that they can deliver a solution in-house, is raising concerns and increases the risks of failure. Especially when the support from senior management and multiple functions of the organisation are hesitant. Proposing an in-house MDM solution to people that their main concern is how to increase their revenue and how to battle the current economic circumstances, shows a lack of maturity and lack of risk awareness.
5. **Timing**. Timing is one of the most crucial elements when an MDM implementation is considered. It is even more important when an organisation initiates a BTP. The results from both audit reports and the interviews shown that data definitions and data mapping to multiple disparate data sources was the most time-consuming element of their projects. Also, the results presented that it is highly recommended that a new ERP implementation and a BTP overall should be aligned with a data strategy. Especially, a BTP that includes an MDM initiative, and a data governance policy is under substantial risk based on when the data strategy implementation occurs. If the processes are not aligned with the data, then the definition of these processes will take longer to be mapped. As a result, the budget and the costs of the programme are going to differ considerably from the original plans. This

can result in BTP failures due to lack of resources (financial and task force). On Chapter 2, there is a detailed proposed MDM roadmap as suggested by Infotech in March 2014. This framework should be followed and implemented before any major BTP that contains an MDM and data governance policy.

The table below summarises the findings from the audits and interviews.

Highlights	Interview findings
Data Quality Challenges	• People willing to invest for improved Data Quality • Data Quality is important to every business function • The longer it takes to resolve Data Quality, the more it costs to the business • The costs are external (fines, penalties etc) and internal (additional tasks for existing resources or additional resources to resolve existing issues) • Delay on process improvement
Change Management	• Communications between different business functions is challenging • Important changes in business processes are not published to other departments. • General assumption that what is clear for a job role should be also clear to someone else in a different role. • Secrecy and reluctance to share
Senior Management Commitment	• Overall support from the senior management on the theoretical basis • Misalignment between senior management and “doers” on the level of support that they give and get • Not enough trust between all parties • Too much time spent on trying to convince Senior Management
Preferred Solution	• Off-the-self solution • Each organisation should focus on its prime line of business and not trying to re-invent the wheel by trying to invest in MDM Research and Development • Showcase a valid roadmap based on the current organisational environment.
Timing	• Data Definitions and Data Mappings is the most time-consuming element, • BTP and Data Strategy should be highly aligned • MDM and Data Governance should be part of Data Strategy which should be part of BTP

Table 15. Key Findings from Interviews.

5.5. Audit Reports and Interview results in comparison to the proposed MDM framework.

Based on the interview results and the audits, none of the organisations of the participants was following a specific framework when they started their projects. The audit reports revealed that Company One was trying to recover from a failed ERP implementation that cost to the business a significant amount of money. Company Two was at the beginning of BTP and they wanted to make sure that there is a data strategy in place to support this BTP. However, Company Two had already planned the BTP without taking into consideration any data element.

MDM and data quality were the crucial elements in both cases. Also, as part of the MDM process, the initiation of the data governance process should be defined and presented to the business.

The initial part of any MDM and data governance process as described by the framework is to have a very well-defined roadmap. Based on the results, there was no roadmap or plan on how the master data would be managed and how the data quality would be ensured. There was a desire that the data quality should be the best possible. That is why both organisations but also the participants of the interviews were involved in projects that had a lot of investment in technology.

This investment, however, required to have all the involved people aligned to the same goal. Based on that, and based on the framework, the case of the MDM should be made clear to the appropriate sponsors, to understand the benefits of their investment.

At this point, almost all the participants and the organisations were facing multiple challenges. Making the case for an MDM project and data governance policy especially after a failed ERP implementation or just before a BTP requires effort. It is always difficult to showcase the benefits that a solution such as MDM could bring to the organisation, especially when the trust between senior management and the team responsible to deliver is not there.

During the interviews, there were many instances that the participants were suggesting that it is very challenging to convince the sponsors of what it is required to be achieved. As they suggested, there was more time spent on convincing than delivering. This would never assist in making the case for the solution. However, apart from convincing, there was a challenging relationship between the teams. The BTP team was suggesting that the new systems would be the solution to the problems that the organisation was facing up to that moment. Their belief

that the new systems would be a completely autonomous solution that could overcome any data related issues (without any proof) was not strengthening the case of the MDM. The most significant part of the budget is always spent on systems that are operated by most of the end-users. MDM solution requires a specific group of people that they would be the operators. The main functionality was not visible to the biggest part of the organisation; therefore, it was difficult to make the senior management to understand the benefits.

In summary, making the case for an MDM solution is always challenging when the MDM process competes with an ERP project even though everybody wants to achieve the best possible outcome.

Except for making the case for MDM, the second biggest challenge that people were facing was setting the expectations. Senior management should not have to worry or think about how to complicate an environment (from a technical perspective) can be. There is little to no interest in how many data sources have to be used and how many integrations need to be developed.

Since there is a belief that the MDM solution can solve the data quality challenges. They expect that the process will perform as it should. However, setting the expectation is very important as most of the people believe that after the installation of the software, all the problems have been resolved. A maturity assessment should determine the state of the data within the organisation, and based on the results, the expectations should be set accordingly.
Also, the implementation style as described in chapter 3.4. must be very detailed during its definition because this would set the expectations on how the process operates. It is very important to define the phases of the process and be able to measure the outcome of each phase. By following that way, the sponsors could understand the behaviour of the MDM process and understand the reasons why a data governance policy would be required. Based on that, the data governance process could be amended and modified based on the new releases of the MDM solution.

Based on the comment above, the evolution of data governance would be able to be applied in a more controlled environment. The results have shown the communication between departments is very challenging. People are not fully aware of the process or they are not willing to change the process that they are currently used to. The less invasive the implementation of

the data governance is, the highest success rate its implementation will get. Also, the end-users will feedback to senior management the level of satisfaction that will get from the data governance process. Based on the feedback, the senior management will be able to assess the policy and support it when more rules are introduced.

Overall, all the people that are involved in an MDM process and the introduction of a data governance policy should have their objectives aligned and support each other's tasks. If this is achieved, the implementation of the MDM process should be more achievable.

The common outcome from the results was that all the participants and the departments within the organisations were not aligned. Everybody had their own goals, and everybody was trying to achieve and meet particular objectives in an individual or departmental level ignoring the overall benefits for the organisation. If senior management could not support the process in practice, the framework could not work. Also, if the participants could not align their objectives but also understand that any change would benefit the business in general, the framework could not work either. And finally, if the participants and the senior management could not understand that both MDM and business transformation are part of the same process, the framework could not work.

The following table represents a summary of the key findings.

Highlight	Company One	Company Two	Interviews
Framework	No Specific Framework	No Specific Framework	No specific Framework

Environments at the time of the Research	Recovery from a Failed ERP Implementation which cost a lot of money	Beginning of BTP. No specific Data Strategy in place	During Phase Two of a BTP.
Data Quality	High Importance		
Roadmap	No Roadmap on how to manage Master Data and ensure Data Quality. But the desire for Data Quality		
Investment to achieve a successful BTP and Data Quality	Investment in technology to achieve.		
Framework: Making the Case for MDM	After a failed ERP, it was challenging to convince Senior Management. Trust between parties in jeopardy.	Challenging relationships between teams.	Challenging in convincing Senior Management. More time spent convincing than delivering.
	Suggestions that new systems would resolve all issues and that they would be completely autonomous that would overcome and data quality challenges.		
	MDM is not a system that is used by the majority of End-Users so it requires more internal promotion.		
	Setting correct expectations	Expectations aligned with the current situation	Expectations and understanding were varied per role
Framework: Preparations for MDM & Maturity Assessment	Believes that the only thing that is required on the current state is a technology that can support the people and the process. However, there isn't a process that can tell people what to do.	No Data Strategy. However, that is why they had the Audit, however, the programme in specific areas was proceeding without taking into consideration any data related constraints.	No maturity assessment. In some cases, a programme director was suggesting that the data take on the process should be ignored if the integration process was too difficult to be achieved.
Framework: Implementation style	Unrealistic targets for achieving an MDM implementation style.	Realistic on what it can be achieved based on time, budget and future process definition.	Falsely belief that ERP system would perform MDM as an out of the box functionality
Data Governance	Challenging communication. The process is not clear among users and functions. Users believe that applying a data governance policy is invasive.		
End Goal	The not aligned common goal between participants and departments within the organisations. Everybody had their own goals with particular objectives ignoring the end goal. Should be clear that MDM and BTP are part of the same solution and investment		

Table 16. Key Findings cross-case.

5.6. Contributions based on the findings

The main research question is as follows:

"***How may the invasive circumstances during a Business Transformation programme be addressed to enhance decision-making with its impact on Master Data Management?"***

So far, most studies on MDM have focused on the technical elements of the implementation. In this study, the emphasis was on the environment that such an implementation takes place in. Identification of the level of support that the implementation team receives from the senior management and how the interaction between multiple functions and departments affect the progress. Identification of how the expectations are set and what is the feasibility of these expectations to be met. These are all invasive circumstances during a BTP that affect decision-making and have an impact on the MDM.

Based on the theoretical foundation, the interviews and the audits, all the research objectives have been met. The way that these objectives have been met, assisted to develop an answer to the main research question. The answer to the main research question can be divided into three major interrelated decision topics which are aligned with the research objectives as stated in section 1.3. These decision topics are as follows:

- **Decision Topic 1**. Invasive circumstances that can affect the MDM and consequently the BTP.
- **Decision Topic 2**. Key points of the MDM Framework and Roadmap that can jeopardise the MDM and the BTP.
- **Decision Topic 3**. Challenges created by timing that affect the MDM and consequently the BTP.

Decision Topic 1

The first topic is divided into 4 different parts.

The first part is the pre-existing Data Quality challenges. These challenges are a cost to the business which is related to the operational data quality challenges. The business will have to allocate resources to address these data quality challenges. As a result, existing tasks cannot be done because people are occupied to on resolving these challenges; as a result, other data and operational activities are not progressing.

The second part is change management and control. The challenge appears when there is not an established communication framework or there is one, but it is poorly managed. This is a result of specific business process changes takes place in silos by people that operate these changes in secrecy, or they do not adapt to these changes because they believe that they are not part of their usual activities. This is due to a lack of understanding of what is mandatory and what is not.

The third part is the commitment from the senior management. The invasive circumstances appear when there is not enough support of the cause or there is not a clear and deep understanding of what is needed to be achieved because it has been explained in very technical terms to a non-technical audience. That usually causes also misalignment between supporters and implementers on the support levels. This misalignment leads to low levels of trust between the teams. Which leads to more time spent on convincing of what needs to be done instead of making progress.

The fourth and last part of this topic is the selection of the preferred solution. The results have shown that in some instances there was not enough understanding of the difference between off-the-self and a built-in house solution. There was a misjudgement of the appropriate solution for the size of the organisation, and there was an underestimation of the required operational and capital budget. There was also a misjudgement on the resources with the right skills for the programme and an underestimation of the existing skills and knowledge on the internal or external resources by overvaluing or undervaluing the programme. Setting expectations is also important and it is essential to be realistic with yourself and to others, otherwise, there is a possibility to lose the direction of what needs to be achieved and to have the misconception that the solution will solve all the data and process challenges on its own. This is not the case. It needs to be a balanced alignment on BTP and MDM of People, Process and Technology.

Table representation of the topic is on the next page, followed by a graphical representation.

Topic Level	Reasons that can affect the decisions
T1	Invasive circumstances that can affect the MDM and consequently the BTP

T1.1	Pre-Existing Data Quality challenges
T1.1.1	Business Cost due to operational Data Quality Challenges
T1.1.2	Internally allocated Resources to address these Data Quality challenges
T1.1.2.1	Existing tasks cannot be done because people are occupied
T1.1.2.2	Other data and operational activities are not progressing
T1.2	Change management and control
T1.2.1	Communication framework non-existing or poorly managed
T1.2.1.1	Business Process change is Silos
T1.2.1.1.1	People operate under secrecy
T1.2.1.1.2	People do not adapt to changes because they are not part of their usual activities
T1.2.1.1.2.1	Lack of understanding of what is mandatory and what not
T1.3	Senior Management Commitment
T1.3.1	Not Enough support of the cause
T1.3.1.1	Not clear and deep understanding
T1.3.1.1.1	Too Technical
T1.3.1.2	Misalignment between Supporters and implementers on the level of support
T1.3.2	Not enough trust between teams
T1.3.3	More time spent on convincing the senior management of what needs to be done instead of making progress
T1.4	Preferred Solution
T1.4.1	Not enough understanding of the difference between off the self and build in house solutions
T1.4.1.1	Misjudging the appropriate solution for the size of the organisation
T1.4.1.2	Underestimate of the required Operational and Capital Budget
T1.4.1.3	Under resource or Over resource the Programme with the right skills
T1.4.1.4	Underestimate existing skills and knowledge on internal or external resources by overvaluing or undervaluing it
T1.4.1.5	Not Being realistic on your expectations but also setting expectations to others
T1.4.1.6	Losing Direction of what you want to achieve
T1.4.2	A misconception that the solution will solve all the data and process challenges on its own
T1.4.2.1	Unbalanced alignment on BTP and MDM of
T1.4.2.1.1	People
T1.4.2.1.2	Process
T1.4.2.1.3	Technology

Table 17. Research Question, Answer Topic 1. Invasive circumstances that can affect the MDM and consequently the BTP.

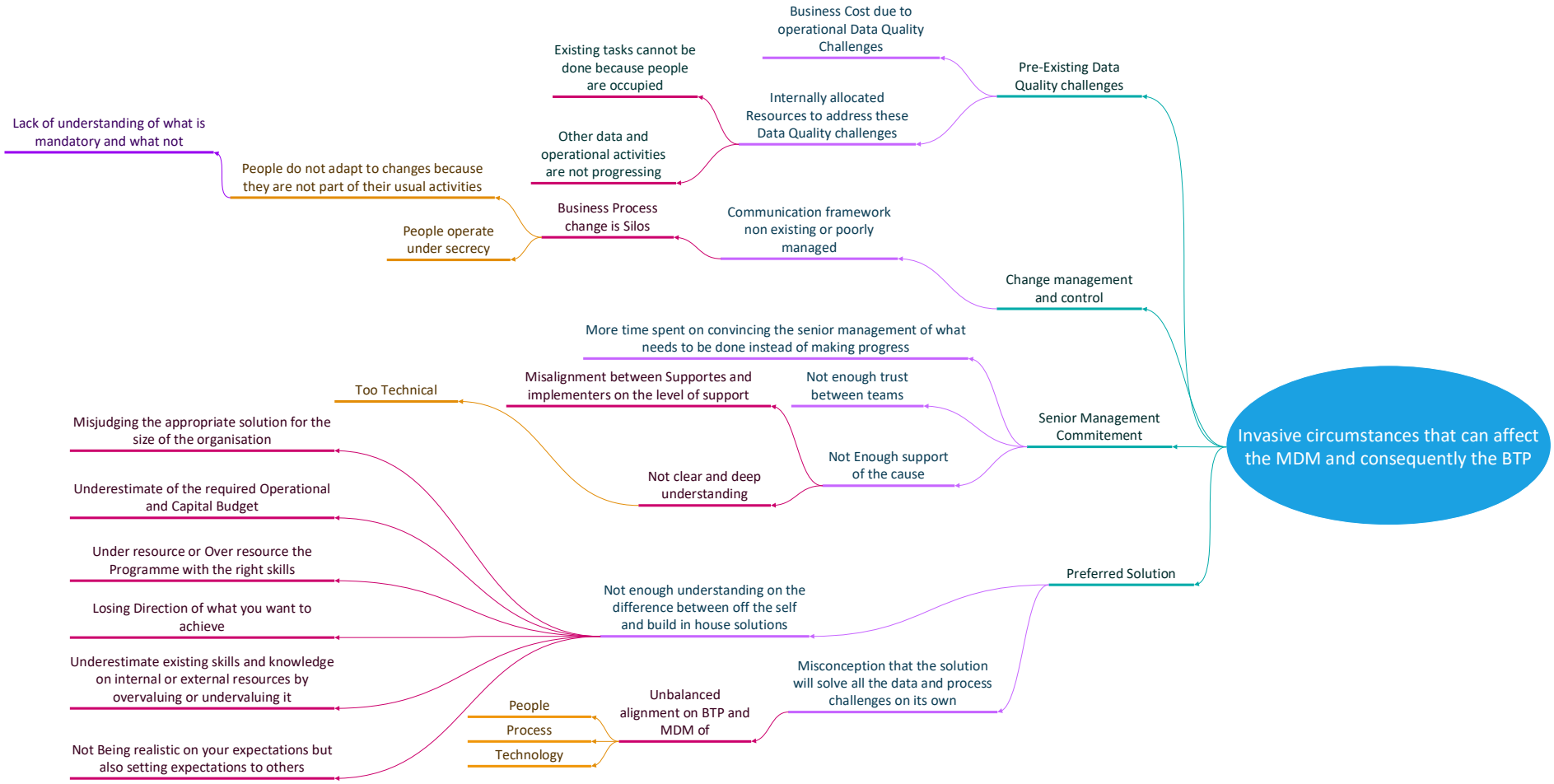


Figure 47 Topic 1. Invasive Circumstances that can affect the MDM and consequently the BTP

Decision Topic 2

The Second topic is divided into 5 different parts.

The first part is the Data Maturity Assessment that is defined on the MDM framework and when this takes place. The challenging situation appears when the Data Maturity Assessment happens after Phase 1 of the BTP. If this is the case, it means that there is no clear understanding of timescales based on the existing resources but also in case that extra recruitment is required for more skilled personnel. If there is a requirement for more recruitment, then there is additional time to be added to the plan until the necessary people join the programme. Besides, any recruits, they will need time to adjust and familiarise themselves with the company's environment. If there is a decision to upskill and educate existing resources; that would require additional time on the planning as well. All the resources (existing and recruits to support the programme) would need to be familiar with the existing and the future environment. Additional time should be assigned and evaluated based on the process changes that the new environment demands as well as the actual resources that would be needed to resolve any data challenges.

The second and the third part is the MDM Roadmap and how making the case for MDM can affect the programme. It is challenging to present a well-defined roadmap when there are uncertainties like all the above. As it has been described on chapter 2, it is necessary that when making the Case for MDM be in a position to defend any claims but also to be setting the expectations in a way that the audience can see understandable results. The language should always be appropriate for the right audience and it should be comprehensive for the IT users that will have to understand the business in a technical level and at the same time not to be overly technical for the business users that they do not have the technical capability and need to be in a position to understand the technicalities and the technical terms.

The fourth part is the senior management confidence. When there is no progress presented to the senior management, the confidence levels will start dropping. That is why there should be specific pre-defined Targets that will have to be achieved by a specific set of dates followed by specific targets and KPIs. There should also be a predefined set of dates for progress updates, so every participants' expectations are met.

The fifth part is Investment in Technology. Systems, applications and processes within this ecosystem generate data and consume data. Data Quality defines the ROI of the investment on all the channels of business and results in revenue increase. The revenue increase can secure investment in more people and overall, more investment since the confidence levels are arising. However, technology on its own cannot ensure Data Quality. Specific rules and policies are coming out of data Governance and as long as people comply with these regulations there is a better success rate on ensuring data quality. Following the above, technology can be configured to follow specific rules and apply specific constraints on the processes. What technology cannot do is to be a part of internal clashes and internal politics. Internal politics prevent results and consume time to irrelevant tasks to both MDM and BTP and can distract entire programmes from the overall goal. Every team has a set of tasks to complete and to complete these tasks there needs to be a focused effort. Teams should focus on the common goal and not get consumed on un-necessary competitions between them. Every team has a set of tasks and every task is equally important with any other task. In addition to that. Everybody has to work as a team in collaboration. Nothing can be delivered by one person only. Everything requires a team effort and there where the focus should be.

Table representation of the topic is on the next page, followed by a graphical representation.

Table 18. Research Question, Answer Topic 2. Key points of the MDM Framework and Roadmap, that can jeopardise the MDM and BTP (Next Page)

Topic Level	Reasons that can affect the decisions
T2.1	Data Maturity Assessment
T2.1.1	Data Maturity Assessment happens after Phase 1 of the BTP
T2.1.1.1	No clear understanding of the following
T2.1.1.1.1	Timescales evaluation based on existing resources
T2.1.1.1.1.1	In case of a Requirement for Extra skilled Resources
T2.1.1.1.1.1.1	Recruitment Time
T2.1.1.1.1.1.2	Educate & Upskill existing Technical Resources
T2.1.1.1.1.1.3	Familiarisation for Resources with the company's new environment
T2.1.1.1.1.1.3.1	Existing & New Environment
T2.1.1.1.1.1.3.2	Existing & New Resources
T2.1.1.1.2	Timescales evaluation based on Process Changes
T2.1.1.1.3	Resources required to resolve and Data Challenges
T2.2	MDM Roadmap
T2.2.1	Fail to Present a well-defined Roadmap
T2.3	Making the Case of MDM
T2.3.1	Not Choosing the right language
T2.3.1.1	Not enough details for the IT Users
T2.3.1.2	Too Technical for the Business & End-users
T2.4	Senior Management Confidence
T2.4.1	No Progress presented to Senior Management
T2.4.1.1	Not Pre-Defined Targets
T2.4.1.2	No KPIs
T2.4.2	No Pre-Set Dates for Progress Updates
T2.5	Invest in Technology
T2.5.1	Technology Generates Data
T2.5.2	Technology Consumes Data
T2.5.2.1	Data Quality Defines the ROI of the investment
T2.5.2.1.1	Senior Management Understands Results
T2.5.2.1.1.1	Results Lead to Revenue
T2.5.2.1.1.2	Results from Secure Investment in People
T2.5.2.1.1.3	Results Secure More investment
T2.5.2.1.1.4	Results Secure Senior Management Commitment
T2.5.3	Technology On its own does not ensure Data Quality
T2.5.3.1	Compliance with Data Governance Ensures Data Quality
T2.5.4	Technology Does Not understand internal clashes and internal politics
T2.5.4.1	Internal Politics prevents Results
T2.5.4.1.1	Effort and time spent on irrelevant tasks to both MDM and BTP
T2.5.4.1.2	Distracted from the overall Goal
T2.5.4.2	Every team has a set of tasks to complete
T2.5.4.2.1	Competition between teams prevents focus on the common goal
T2.5.4.2.2	None of one team's tasks is more important than other teams' tasks. Every task has its purpose
T2.5.4.3	Nothing can be delivered by one person only
T2.5.4.3.1	It is a team effort that delivers and not an individual's short moment of glory

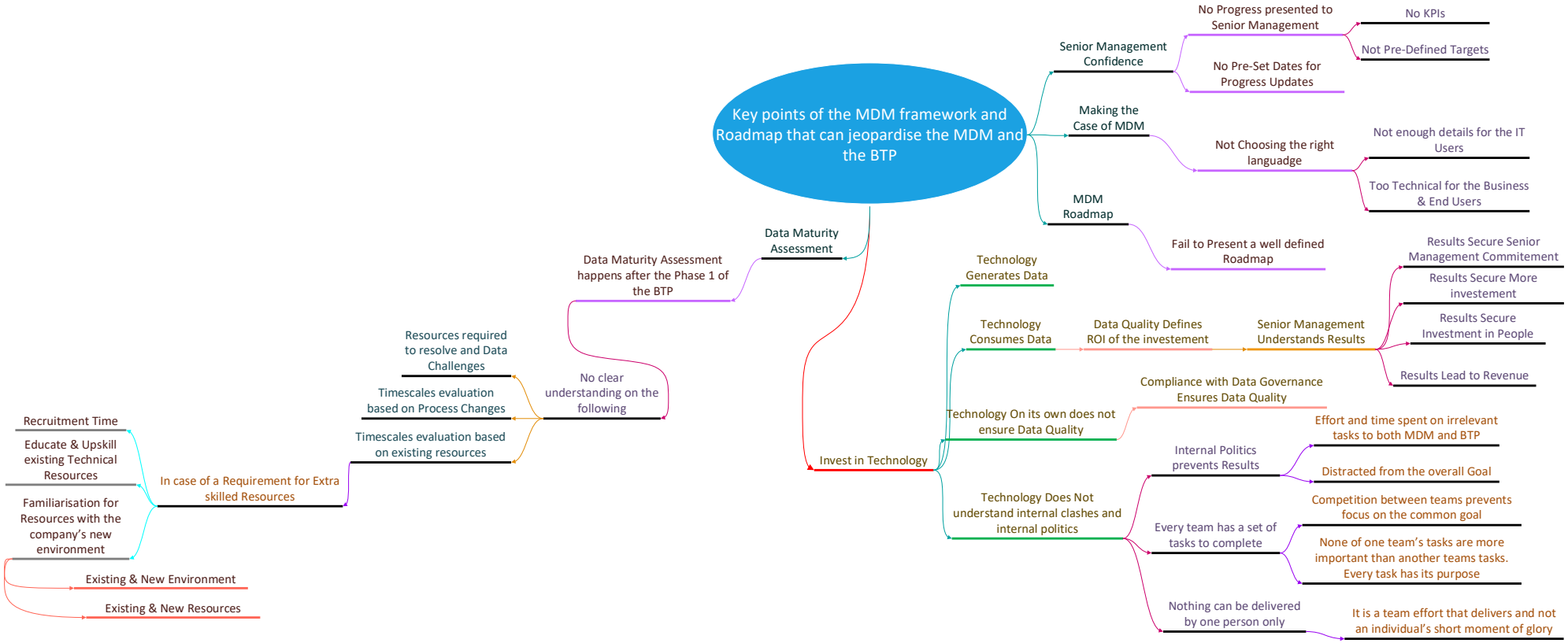


Figure 48 Topic 2 Key points of the MDM framework and Roadmap that can jeopardise the MDM and the BTP

Decision Topic 3

The third topic is divided into three parts and it is related to Objective 3.

This topic is focused on the timing of the events and what challenges are created by choosing the right time to perform specific tasks. The first part is focused on the Data flow and the timing is critical on when and how the flow is designed. The process flow should align with the Data flow and the data flow should serve the new process flow. The challenges on that as it revealed on the interviews and the audits is that there is a lack of understanding of the existing data locations, existing data rules, data constraints and data dependencies. There were also instances that there was a lack of understanding of the new applications' ecosystem, data requirements and the rules of the new application. These rules should be fulfilled by the Data Governance policy which should be introduced as part of the new application ecosystem in a non-invasive way by educating the users regarding the reasons for the change on every channel of the business.

The second part is the Data strategy. As per Objective 3, Data Strategy should not be initiated as a process after phase one or during phase two of a BTP and Data Strategy should be defined by the Data Governance during a BTP. Data Strategy should define the data governance for the implementation phase of BTP and after the programme's life. The strategy should take into consideration any previous rules and polices but should not be limited to previous facilities within the process.

The third and final part focuses on data definitions and mappings. As per Objective 3, these two tasks require more effort and time that is why it is important to build more contingency and allocation more time in the programme for these tasks along with more technically skilled resources and more resources with the process knowledge.

Table representation of the topic is on the next page, followed by a graphical representation.

Topic Level	Reasons that can affect the decisions
T3	Challenges created by timing that affect the MDM and consequently the BTP

T3.1	Data & Process Flow
T3.1.1	Process flow Should align with the Data Flow
T3.1.1.1	Lack of understanding of the existing Data Locations
T3.1.1.2	Lack of understanding of the existing Data Rules
T3.1.1.3	Lack of understanding of the existing Data Constraints
T3.1.1.4	Lack of understanding of the existing Data Dependancies
T3.1.2	Data Flow Should Serve the new Process flow
T3.1.2.1	Lack of understanding of new applications' ecosystem
T3.1.2.2	Lack of understanding of new applications' rules
T3.1.2.2.1	Apply Data Governance as part of the new applications ecosystem
T3.1.2.2.2	Apply Data Governance in a non-invasive way
T3.1.2.3	Lack of understanding of new applications' data Requirements
T3.1.2.3.1	Educate the Users about the Reasons for and of the change
T3.1.2.3.1.1	End-users
T3.1.2.3.1.2	IT Users
T3.1.2.3.1.3	Business Users
T3.1.2.3.2	Explain in Detail why this process change is vital for the business
T3.1.2.4	Lack of understanding of new applications' Data Integrations
T3.2	Data Strategy
T3.2.1	Data Strategy initiated as a process after the Phase One and during Phase Two of a BTP
T3.2.2	Data Strategy defined by Data Governance during BTP
T3.2.2.1	Data Strategy Should define Data Governance for the following
T3.2.2.1.1	Implementation Phase of BTP
T3.2.2.1.2	After BTP's life
T3.2.2.1.3	Taking into consideration any previous rules and polices but not limited to previous facilities within the process
T3.3	Data Definitions and Data Mapping
T3.3.1	Not enough contingency allocated on extra time
T3.3.2	Not enough technically skilled resources
T3.3.3	Not enough resources with the process knowledge

Table 19. Research Question, Answer Topic 3. Challenges created by timing that affect the MDM and consequently the BTP

Figure 49 Topic 3 Challenges created by timing that affect the MDM and consequently the BTP

Following the presentation of the three different topics, the next figure represents all the topics interconnected under one diagram.

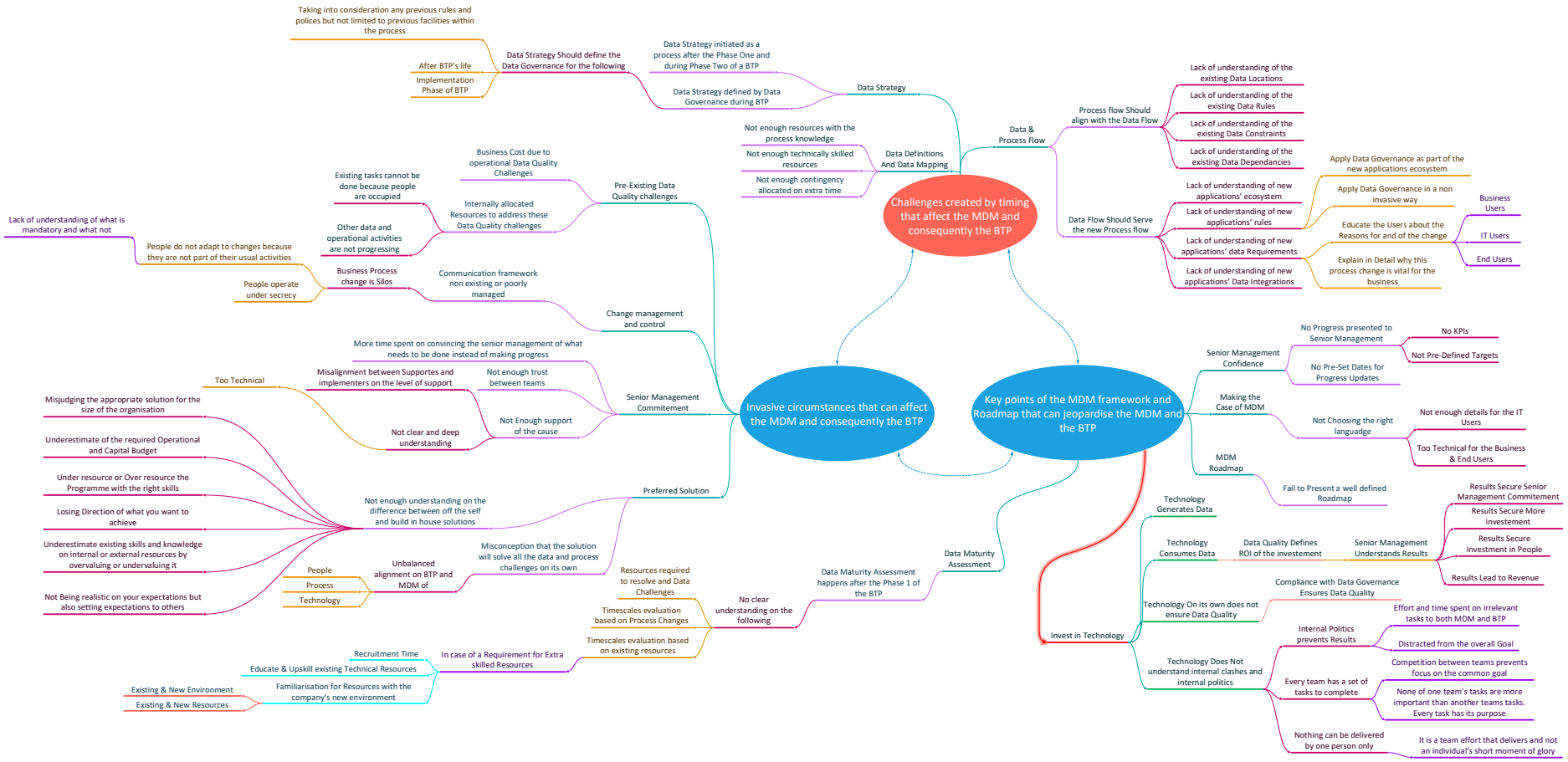


Figure 50 Overall Answer to Research Question

Overall Summary of the Decision topics

All three topics are related and interconnected. Invasive circumstances can appear at any point during any phase of a BTP or an MDM process with a specific framework and roadmap. Challenging data quality situations can appear at any point without taking into consideration any change management controls no matter what the system solution is; the senior management can increase or decrease the level of confidence to the project at any point affecting the much-needed support and commitment.

All three topics have been derived from the research objectives and each objective is aligned with more than one topic.

- Topic 1 has been formed based on the outcome of the research objectives 1,4 and 5.
- Topic 2 has been formed based on the outcome of the research objectives 2,4, and 5.
- Topic 3 has been formed based on the outcome of the research objectives 3 and 6.

All three topics together propose an MDM-impacted BTP decision model. This model offers an enhanced understanding of the reasons behind the decisions during a BTP concerning MDM, and how these decisions consequently affect the successful implementation of a BTP.

Chapter 6: Conclusions and Further Work

This research aims to establish the importance of the MDM initiative during a BTP and to define the decisions and the circumstances that these decisions were made during the programme can affect its successful implementation. This is valuable to any future MDM initiative as well as any BTP.

This chapter concludes to what extent the aim and research objectives have been addressed and how the main research question has been answered. Contributions to the body of knowledge are described along with identifying further areas of research in this field.

6.1. Contributions to knowledge

The primary contributions of this research are as follows. These are aligned with the research objectives as stated in section 1.3.1.

1. *Provision of focus through the acquisition of knowledge of the reasons that cause data quality problems within an organisation*.

All the organisations and all the participants are aware that there are multiple **data quality challenges** within their business processes. That is why there is a level of investment to address these challenges and find a resolution. Data quality is important for every function within the business and the longest period a challenge is not resolved, the highest the cost for the organisation overall would be. As it was presented from the interviews and the audits, the cost is not only measured on the number of fines or penalties a company could get from a vendor or a customer. There are costs for every resource that is allocated internally to resolve these challenges. People should disengage from their normal activities, so they can be allocated to a resolution of a data quality issue. As a result, other activities are not progressing in the originally estimated time. Based on that, new budgets should be reconsidered and allocated, and the cost is multiplied across all the functions.

As it has been described in detail in chapter 2.3.1., the processes that are mainly responsible for these data quality challenges and disrupt the information are divided into three different groups: the processes that bring data from outside the database, the processes that use data that are handled inside the database, and the data that already exist in the database but becomes obsolete mainly due to time, and not due to changes in the database.

This objective contributed to the formation of decision topic 1 by focusing on the Data Quality and the challenging circumstances that may appear during an MDM and consequently during a BTP. Figure 47 shows in detail the reason for Data Quality Challenges.

2. *To focus on the existing enterprise data quality and data governance methodology that is part of a BTP and the extent that is taken into consideration by the people that are involved in the programme.*

Based on the interview results and the audits, none of the organisations of the participants was following a specific framework when they started their projects. MDM and data quality were the crucial elements in both cases. Also, as part of the MDM process, the initiation of the data governance process should be defined and presented to the business. The initial part of any MDM and data governance process as described by the framework is to have a very well-defined roadmap. Based on the results, there was no roadmap or plan on how the master data would be managed and how the data quality would be ensured. There was a desire that the data quality should be the best possible. That is why both organisations but also the participants of the interviews were involved in projects that had a lot of investment in technology.

At this point, almost all the participants and the organisations were facing multiple challenges. Making the case for an MDM project and data governance policy especially after a failed ERP implementation or just before a BTP requires effort. It is always difficult to showcase the benefits that a solution such as MDM could bring to the organisation, especially when the trust between senior management and the team responsible to deliver is not there. However, apart from convincing, there was a challenging relationship between the teams. The BTP team was suggesting that the new systems would be the solution to the problems that the organisation

was facing up to that moment. Their belief that the new systems would be a completely autonomous solution that could overcome any data related issues (without any proof) was not strengthening the case of the MDM.

The common outcome from the results was that all the participants and the departments within the organisations were not aligned. Everybody had their own goals, and everybody was trying to achieve and meet particular objectives in an individual or departmental level ignoring the overall benefits for the organisation. If senior management could not support the process in practice, the framework could not work. Also, if the participants could not align their objectives but also understand that any change would benefit the business in general, the framework could not work either. And finally, if the participants and the senior management could not understand that both MDM and business transformation are part of the same process, the framework could not work.

This objective contributed to the formation of decision topic 2 by focusing on the critical areas that can jeopardise the BTP.

3. *Provision of focus through the knowledge representation of the MDM initiative and clarify the support to specific design areas during the planning of the MDM.*

Based on the interviews, the audits and the theoretical foundations, the critical subject that needs to be taken into consideration as a priority, is the timing of the Data Strategy since MDM and its architecture is (or it should be) part of the Data Strategy. When planning a BTP, the Data Strategy should be treated of equal importance and should be started at the same time as the BTP, if there was no one in place already. If there was already a Data Strategy initiative within the organisation, the strategic objectives should define what needs to be done in case of process and technology change by the people who run it.

Based on the above, there are some constraints regarding Data Governance and the priorities that specific activities take place. Data Strategy should be defined by the Data Governance during the BTP because then the Data Governance is limited only on the requirements of the BTP and not the overall application and Data Ecosystem. For that reason, Data Strategy should define the Data Governance process for the implementation phase of the

BTP taking into consideration the full application ecosystem on the existing state, the implementation state and the future state. Data Strategy should take into consideration any previous rules and constraints but should be limited to any previous facilities, applications and data flow.

Data Integration is part of every Data Strategy and consequently part of BTP. The most vital part of any data integration is the data definitions of the existing data systems and the data mapping to the new. Both interviews and the audits show that these two parts were the most challenging. That is why on the design phase of the MDM more time should be assigned to defining and mapping the data on multiple interconnected exercises. What was also challenging was the fact that there were not enough skilled resources and not enough resources who understand the process.

By succeeding on the data definitions and the data mappings, there is a better understanding of the existing data locations, existing data rules, existing data constraints and existing data dependencies; with this level of understanding, the process flow will align with the data flow.

Also, it is more efficient to apply data governance as part of the initial release of the new applications' ecosystem to the end-users, business users and IT users. In that way, the transition to the new rules, data requirements and integrations becomes non-invasive since there is a better understanding of the reasons for change and its importance to the business process. By that way, the data flow should serve the new process flow.

In conclusion, to achieve all the above, the MDM framework should run before any BTP to the point of Data Maturity assessment. By that way, there will be a full understanding of what resources are required to tackle any data challenges, how much time will be required based on the process change and how much time will be required based on the existing resources. In case of any requirements for additional skilled resources, there will be a more accurate plan in place based on how long will take to recruit these resources, how long will take to educate and upskill the existing technical resources and how long will take to familiarise the resources with the company's existing and new environment.

This objective contributed to the formation of decision topic 3 by focusing on how critical the timing and the specific sequence of events are during a BTP.

4. *To enable knowledge sharing and reuse to practitioners to deepen the understanding of the challenges that may have to overcome during a BTP.*

The audits and the interview results revealed that in some instances there is a need for more training in matters of Data Management. People responsible for making decisions, focus on matters for their business but not in the recommended sequence. During a BTP where the businesses review their processes and are prepared to change the way they operate, they focus on delivering the process first without considering the data that forms these processes. Another example that highlights the need for more training and education is that people are confused on what master data is and they spend money, time, and effort on creating the content instead of managing the content as per the Company One audit results.

In addition to the need for more training and overall education, people often feel challenged, with unpredictable reactions. The research highlighted two different categories of people:

- The first category is that they are hesitant to change. Changing the way that someone works by applying rules that were not there before, is always a challenge. They will struggle to adopt this challenge as an improvement but as a disruption to their daily tasks.
- The second category is that they are not confident enough about what they are doing and that their jobs will be at risk. The defensive posture that people were taking when they were being asked questions like “Why are you doing it that way?” was very visible during the conversations. People are hesitant to changes. One example is that during the Company Two audit, there was an instance that specific time had to be spent with the finance team for defining the data transition strategy. The focus of this conversation was the suppliers’ data domain. Most of the suppliers did not have addresses and when a question was asked of “why there are not any addresses on the suppliers”, the finance person answered in a very defensive way: “We don’t need the addresses for the suppliers. It is not required, and I don’t understand why you ask about the addresses?” The system that the finance person was using was planned to be the main master source for suppliers for the new ERP. The user was trying to defend their position instead of focusing on the greater picture that the organisation was going through with the BTP.

A different reason that usually a BTP suffers due to data, is the fact that teams work in silos instead of having a board that aligns their tasks and the focus of the project. As a result, the creation of different rules for the same data domain per function. As an example, Company One from the audits had different departments maintaining different data sets for the same data domain. As a result, during the data consolidation process for the data transition, a considerable amount of data to be out of scope due to lack of communication. This caused a considerable delay to the project due to configurations that they had to take place since the data that have been left out of scope was important for specific processes.

As it is described in chapter 3.6 of the literature review, Sherman (2011) highlighted ten different practices that could jeopardise the BTP. During the research period, every example that Sherman (2011) suggested, could be related to a real event within a project or a programme.

1. The first practice that Sherman (2011) highlighted was the “Buy-in but not commitment”. This statement can be related to the audit results from Company One. Company One is willing to invest and already investing a very large amount of money by allocating resources to a team to manage the master data. This team is 19 people and 1 director. The cost of the team to the business is nearly £500,000 annually and by speaking to these people and the director, the business does not allow this team to work on MDM. Even though their department is called MDM department. The organisation does not support with the appropriate tools and their tasks are limited in creating content for the content creation team using Excel spreadsheets.
2. The second practice is the “Ready, Fire, Aim”. Applies to both Company One and Company Two. The organisation has already invested about £500,000 in people before giving them the tools to do their jobs. Company Two, the problem is different. They also follow the “Ready, Fire, Aim” but in a different way. They have invested in a product and an entire transformation of their business process without having analyzed the data that the new system will be required to operate. The Company Two have a specific deadline to go live with the new system and for more than a year, they were focusing on making the system work with test data that is not related to the operational data that they will need. They have not tried to insert any data from the existing systems into the new environment

because the programme director believes that this will be a task that will require analysis in a later stage.

3. The third practice which is “Trying to solve world hunger or boil the ocean” applies to both Company One, Company Two as well as many of the interview participants. The aspirations are too high for both companies, but the execution could not match their aspirations. For Company One, they have assigned 19 people to manage their master data but on the other hand, they do not have the experience and they do not have the tools to operate this task. Also, they have been tasked with multiple other tasks that consume at least 60% of their time instead of focusing on one thing. In comparison, Company Two is trying to solve their problems based on the functionality of the new solution. They are trying to test it and at the same time, they are trying to define the processes based on test data. All these actions by the same people at the same time. During the interviews, the MDM Director suggested that he can build an in-house MDM solution without any development resources while he has to manage 19 people.
4. The fourth practice which is “The Goldilocks Syndrome” can be identified in the interviews. One of the BTP Directors states that “We are not so interested in taking data on if the data is inconsistent.” That is a “high-level” approach. On the other hand, the complete opposite version of requirements shows a lack of understanding in the process is that Company One has identified as “must-have” requirements for the product domain 1489 attributes.
5. The fifth is committee overload. During the audits, it was common that every meeting that was taking place during the BTP, as well as the MDM process in Company One, was overcrowded. The decision-makers were over-ruled by people that they were not decision-makers. Everyone had to mention the multiple challenges that they were facing. As a result, a decision could not be made due to time constraints and any potential action was postponed for a different meeting. This had, as a result, to have most of the time spent in meetings and not delivering the scheduled tasks.
6. The sixth practice which is “Failure to implement” can be associated with an example from the fifth practice. Too much time is spent in meetings as a result people to be

distracted from what they are supposed to do. Also, there is another element that could be related to this practice for both companies. This was the lack of communication between the teams. People were focusing on a task when this task was completed, they were reporting the completed status, but this status update was not being communicated as a result people were reporting project obstacles on tasks that had already been completed and due to lack of communication they were reporting that they could not proceed with their tasks.

7. The above example also applies to the seventh practice which is “Not Dealing with the change management”. This also applies to the change of process when the MDM practice is in place. In most cases, the business performs a fundamental change in their operations which they think that it is not necessary to inform IT of this change. As a result, the already developed integrations or processes that have been developed by IT, to ignore this alteration on the data that this process may consist.
8. The eighth practice can be related to the interview answers. One of the BTP Directors insisted on many occasions that the new system will manage the data, thus there is no need for an MDM process or a detailed data governance policy. The BTP suggested that all the static data will be managed by the new environment. That is why there was not any requirement for additional budget to be spent on analysis at the time. As it was stated, it was more important to have the system in an operational state than focus on the data.
9. The ninth practise which can be related to Company One is the lack of a sustainable and ongoing process. As described in Chapter 2, the most important part of the framework is to set achievable expectations with quick results. If the business could not get any results within a specific timeframe, the possibilities are that a potential disagreement between departments will arise. Consequently, the sponsors will challenge the feasibility of the project and the confidence level will be under pressure.
10. The tenth practises which can be identified in both Company One and Company Two is the fact that during their BTP they ignored any underling (shadow) systems. This was happening because multiple project teams were working in silos, isolating themselves

from the other business functions. These isolated teams were building an environment that was aligned with their tasks and generate and manage data in their way. The way that the processes were treating this data, should be communicated. In Company One, for example, there was an underlying system that was not part of any other systems that were processing data. This system was analytical, and the owner was exporting data from multiple different sources. After performing manual adjustments and with no validation from any other system, a set of information was distributed to the business, and nobody could understand the way that this data was being created. This system had been left outside the scope because it was not documented, and it was not a part of any architecture design diagrams. However, when there was an attempt to recreate the process in the new environment, it was impossible to generate the required outcomes. As a result, more development and modifications had to be applied to the BTP to make it operational as per the expectations.

This objective contributed to the formation of decision topic 1 and decision topic 2. The knowledge that gained by this objective helped identify multiple invasive circumstances that affected decisions and impacted the BTP in multiple stages and are reflecting on decision topic 1. The knowledge that gained by this objective helped identify critical areas that even when people are trying to follow the theoretical best practices, can still jeopardise the programme based on their programme's progress. These instances are reflecting on decision topic 2.

5. *Provision of an understanding of the knowledge which is derived from the examination on how professionals in different roles respond to the importance of Master Data and Data Governance during the Business Transformation Programme and also, provision of an understanding on how BTP is affected by their perspective.*

All the participants agreed that data quality is important and vital for their projects and the business in general. Each one of them for their reasons. However, in most cases, the project managers and the project directors were prioritising the delivery of their respective projects based on the deliverables. To some extent, meeting the delivery schedule was more important

than the data transition from the old systems to the new. The business analysts were more focused on analysing and designing existing and future processes assuming that the data quality should be a concern for a different department.

All the participants agreed that managing their master data is very important for all their projects and their businesses in general. However, every group of participants had a different focus. All the participants agreed with the current state of managing the master data, and all of them suggested that there was not a properly set up process for this type of activity.

With regards to data strategy, all the participants had different views and understanding since each function has different objectives. There are different views regarding the **support** that data strategy receives from the **senior management**. All the other participants agreed that the business process knowledge and understanding is not always completed. The missing element, no matter the size, is always the most challenging part.

Similar to the previous objective, this objective also contributed to the formation of decision topic 1 and decision topic 2. The knowledge that gained by this objective helped identify multiple invasive circumstances that affected decisions and impacted the BTP this time from a people's and role's perspective and are reflecting on decision topic 1. The knowledge that gained by this objective helped identify critical areas that even when people are trying to follow the theoretical best practices, can still jeopardise the programme based on their programme's progress. These instances which are more focused on how the senior management reacts on potential challenges during a MDM and BTP are reflecting on decision topic 2.

6. *To provide an understanding of knowledge derived from how to approach people from different business functions during the BTP and how their commitment to the data governance framework can affect the BTP and the business, overall.*

All the participants were focusing only on their BTP and the new ERP and they were ignoring any (potentially important) other systems and the integration between the existing environment and the upcoming. There was also an assumption that the new ERP system would be able to manage all the master data as part of the default functionality. As a result, there was not any focus on the data governance

Overall, there was a positive understanding of the importance of MDM from most of the participants. The important outcome is that almost every participant understands the challenges that they currently face within their business and their project regarding the MDM in general. And the most positive outcome is that almost everybody optimistic about the future. This optimism suggests that there are good foundations and willingness within the business for future improvements that are beneficial for the future of MDM and the success of an overall business transformation. The highest risk is the fact that the BTP Directors are principally focused on the process and the deliverables but not on the data. Not enough support from the senior management and that all the teams are isolated when it comes to master data. This is also a risk since based on the conversations, there is not a complete picture of the overall data map and both his team and the project teams work in isolation. That means that there is no coordination and schedule.

All the participants from all the different groups understand the need for data quality. However, there is not an activity in place to ensure that data quality can be ensured and guaranteed due to the excessive use of manually maintained shadow data silos. All these reasons allow the business processes to operate without any control but also any validation. Also, many participants suggested that there is a lack of understanding with regards to business processes but also a lack of reasoning of why a process needs to follow a specific execution. there is a conflict of what senior management is prepared to do, what it is expected to do and is happening. Most of the participants believe that there is not practical support from the senior management and at the same time, senior management believes that the required support is offered. All the participants desire the benefits of quality data, but all of them believe that the responsibility lies within someone else's territory.

Similar to objective 3, this objective contributed to the formation of the decision topic 3 which is the presentation of the challenges that can appear based on the timing that any decision can be made. Each function and each role have a different perspective of time when it comes to tasks as for some business functions specific processes are more important than others hence why they need to be completed based on different priority. The contributions of this objective are reflecting on decision topic 3.

Contributions Summary

All three decision topics and all six objectives are related and interconnected. Invasive circumstances can appear at any point during any phase of a BTP or an MDM process with a specific framework and roadmap. Challenging data quality situations can appear at any point without taking into consideration any change management controls no matter what the system solution is; the senior management can increase or decrease the level of confidence to the project at any point affecting the much-needed support and commitment.

All three topics have been derived from the research objectives and each objective is aligned with more than one topic.

- Topic 1 contributes to the body of Knowledge by identifying some of the invasive circumstances that can affect the MDM and consequently the BTP and their potential impact in multiple stages and it has been formed based on the outcome of the research objectives 1,4 and 5.
- Topic 2 contributes to the body of knowledge by following the theoretical best practices by identifying any potential critical areas that may jeopardise the programme if these areas got overlooked and it has been formed based on the outcome of the research objectives 2,4, and 5.
- Topic 3 contributes to the body of knowledge by examining the merging of the theoretical guideline and the practical implications. The topic identifies that even if the programme directorate follows the best practices, the timing and sequence of executing these tasks may affect the successful implementation and create more invasive circumstances. These findings have formed based on the outcome of the research objectives 3 and 6.

All three topics together propose an MDM-impacted BTP decision model. This model offers an enhanced understanding of the reasons behind the decisions during a BTP concerning MDM, and how these decisions consequently affect the successful implementation of a BTP.

6.2. Thesis Statement

This research aimed to establish the importance of the MDM initiative during a BTP and to define the decisions and the circumstances that these decisions were made during the programme can affect its successful implementation. It has been demonstrated that each Objective has been met and the Research Question has been answered. Therefore, it is concluded that the project aim as has been described in Chapter 1.3 has been met. The contributions to knowledge have been stated in Section 6.2 and the answer to the main research question in section 6.3 of this Chapter.

6.3. Conclusion

This research aimed to establish the importance of the MDM initiative during a BTP and to define how the decisions and the circumstances that these decisions were made during the programme can affect its successful implementation. The conclusion to this research is that the Management of the Master Data and the continuous assurance of the Data Quality is vital for any organisation despite being in a Business Transformation Programme or not. The understanding of the core vital information that resides within each organisation should be available, trustworthy, and understandable by every function of the business.

The qualitative experimental data that has been collected during this study contributed to the body of knowledge by identifying a fracture of the invasive circumstances that can affect any decisions that are due to be made during a BTP with regards to the MDM; and their potential impact in multiple stages of this transformation. This impact can be at the initiation or the investigation for the case of the programme. It can happen during the implementation of the programme or even after the completion of the programme. These circumstances can be grouped to firstly, any pre-existing data quality challenges that had been overlooked for multiple reasons. Lack of communication between the involved parties with regards to changes in business processes or politics. Lack of commitment from the senior management or the sponsors of the project. This lack of commitment could be either time related or financial support. And finally the understanding of what is the most appropriate solution to meet the business need.

Besides, following an in-depth analysis of the theoretical best practices, this study, achieved to identify potentially critical areas that can jeopardise the progress of the programme. The conclusion from the experimental data is that in some cases, even if the programme directorate follows or suggests the best practices, there is always a risk on specific areas to be unsuccessful or challenging. These areas are first, the realisation and the acceptance of how mature with regards to data using the business is. The second area is the interaction with senior management. Any interaction should give confidence to the audience by understanding why this investment is required, and what the roadmap is. Lastly, once all the above have been achieved to a satisfactory level, the selection of the investment in technology is the last challenging decision that can affect every aspect of the programme.

The last conclusion that is generated by this study is the importance of time. It is of strategic significance that specific activities that came out of the experimental data, need to happen at a specific time. Otherwise, the difficulty level for the completion will rise. Any BTP should have set the Data Strategy and every element that this strategy is formed by; along with the BTP and be treated with the same level of significance.

The research in this thesis leads to the conclusion that the above statement is supported. From the Literature, the case study audits and the interviews, the research identified an enhanced understanding of the reasons behind the decisions during a BTP concerning MDM, and how these decisions consequently affect the successful implementation of a BTP. As presented in section 6.1 of this chapter, an MDM-impacted BTP decision model has been formed. This model offers an enhanced understanding of the reasons behind the decisions during a BTP concerning MDM, and how these decisions consequently affect the successful implementation of a BTP.

6.4 Limitations and Further work

This research has led to a better understanding of the domain and resulted in the identification of topics for further work, these are described below.

1. The number of case study audits was limited to two in this research. The available experimental data should be extended to more case studies of challenging circumstances

during a BTP, but also to successful implementation and compare the generated decision model to multiple successful implementations and compare the circumstances that drove these decisions.

2. The number of participants in the interviews was only eighteen. The process of identifying interviewees that meet the criteria for this research type was proven to be challenging. Especially the senior management and any director level interviewees were in some cases non-approachable. More Interviewees that actively participate in a BTP and have an MDM function as part of the Programme's strategy would enable more qualitative data for analysis. In addition to senior management, the role of an MDM director was proven to be challenging since most of the companies do not recruit for this type of role. To have an MDM director participating in the research was a privilege. However, research with more MDM directors as participants would add more valuable qualitative data for analysis since it would be on their interest to affect any decision on the MDM's favour during a BTP.
3. More focus on the change management and the correspondence from the end-users during a BTP. Since a BTP is a transformation and an alteration of the normal day to day activities, there should be an analysis of how the organisation plan to transform the attitudes and behaviour of their employees to react positively to this change by applying psychological methods that translate why people respond in a particular way. What is the purpose to believe in during a BTP, how reinforced this transformation is and what are the incentives for positive reinforcement of the change? Also, how effective is the demonstration of skill that is required for the BTP and what is the perception that the end-users have for the senior management deciding on a BTP and how the senior management can champion the role model of the change consistently.
4. A more detailed analysis should be taken place on the way that Data Governance is applied or planned to be applied during a BTP. As described in chapter 3.7 data governance could become invasive or threatening to the work, people and culture of a business during a BTP. A more quantitative research defining the level of invasiveness and the end-users' reaction could help in understanding in more depth the level of disruption in a BTP and the organisation in general.

References

Aaker D., Kumar V. & Day G., (2004), Marketing Research, eighth edition, Willey, pp. 34 -60

Al-Ruithe, M., Benkhelifa, E. (2017) Analysis and Classification of Barriers and Critical Success Factors for Implementing a Cloud Data Governance Strategy. Procedia Computer Science. 113, 223–232. https://doi.org/10.1016/j.procs.2017.08.352

Al-Ruithe M, Benkhelifa E. & Hameed K., (2017), A systematic literature review of data governance and cloud data governance, Personal and Ubiquitous Computing https://doi.org/10.1007/s00779-017-1104-3

Bajkov A. (2016), Data Governance Worst Practices, http://www.tmcconsulting.ca/advdoc.php?i=201601&a=data-governance-worst-practices.pdf&t, Last accessed: February 2017

Berson A., & Dubov L. (2011), Master Data Management and Data Governance second edition, McGraw Hill Professional. 79-111

Bonnet P. (2008), MAG's Data Enterprise Architecture blueprint: Prebuilt Data Enterprise Architecture for MDM. Orchestra Networks, www.orchestranetworks.com , Last accessed: February 2016

Bonnet P. (2008), MAG's prebuilt data model (business level) A universal prebuilt data model in order to avoid reinventing the wheel. Orchestra Networks, www.orchestranetworks.com , Last accessed: February 2016

Bonnet P. (2008), MDM modelling procedures Master Data Management (MDM). Orchestra Networks, www.orchestranetworks.com , Last accessed: February 2016

Bonnet P. (2009), Case Study: MDM Modelling: How to model the "Address "Business Object when using a Model - driven MDM. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2009), MDM Method Training: Architecture and Modelling Procedures for mastering Master - Reference Data. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2012), Introduction to MDM: Part 1 - A quick IS/IT story Master Data Management. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2012), Introduction to MDM: Part 2 - Data modelling foundation Master Data Management. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2012), Introduction to MDM: Part 3 - Case study - From modelling to UI Master Data Management. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2012), Introduction to MDM: Part 4 - Enterprise Data Architecture Master Data Management. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2012), Introduction to MDM: Part 5 - How to face the data governance issue? Master Data Management. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Bonnet P. (2012), Introduction to MDM: Part 6 - Business object's life cycle modelling Based on the case study: International address management. Orchestra Networks, www.orchestranetworks.com , Last accessed: May 2015

Otto, B., Hüner, K. M., & Österle, H. (2012). Toward a functional reference model for master data quality management. Information Systems and e-Business Management, 10(3), 395–425. DOI 10.1007/s10257-011-0178-0

Dorr B., Murnane R. (2011), Using Data Profiling, Data Quality, and Data Monitoring to Improve Enterprise Information, Software Quality Professional; Sep 2011; 13, 4; ProQuest. 9

Brown J. (2013), Think Big Data, think Amazon Web Services Solution Overview. Amazon Web Services., https://www.business-review-vodcasts.com/amazon-web-services/whitepapers/AWS_BigData_Overview.pdf, Last accessed: June 2014

Bryman A., Bell E., (2003), Business Research Method, Oxford University Methods.

Herrero A., Baruque B., Sedano J., Quintian H., Corchado E. (2015), Advances in Intelligent Systems and Computing 369, International Joint Conference CISIS'15 and ICEUTE'15. DOI 10.1007/978-3-319-19713-5

Holmes A. (2012), Centre for Economics and Business Research Ltd, Data equity Unlocking the value of big data. London: Centre for Economics and Business Research Ltd, http://www.govopps.co.uk/data-equity-unlocking-the-value-of-big-data/, Last accessed: September 2013

Cleven A., & Wortmann F. (2010), Uncovering four strategies to approach Master Data Management, Proceedings of the 43rd Hawaii International Conference on System Sciences. https://ieeexplore.ieee.org/document/5428531/, Last accessed: September 2015

Longman C., (2008), Why Master Data Management is Such a Challenge, in:DM Review, 18, 2008, Nr. 11, S. 18

Collins H. (2010) Creative Research: The Theory and Practice of Research for the Creative Industries AVA Publications, pp. 26-58

DataFlux Corporation, The Data Governance Maturity Model Establishing the People, Policies and Technology That Manage Enterprise Data. DataFlux Corporation, http://www.fstech.co.uk/fst/whitepapers/The_Data_Governance_Maturity_Model.pdf, Last accessed: September 2016

Dreibelbis A., Hechler E. & Milman I. (2008), Enterprise Master Data Management: An SOA Approach to Managing Core Information, Pearson Publishing, pp25-50, 94-101, 132-197

Dudovskiy J. (2018) Interpretivism (interpretivist) Research Philosophy, https://research-methodology.net/research-philosophy/interpretivism/, Last accessed: May 2018

Dyche, J & Levy. E. (2006). Customer data integration: Reaching a single version of the truth. New York: John Wiley.

Eckerson W. (2009), Data Quality and the Bottom Line, Achieving business success through a commitment to High Quality Data, TDWI Report Series. 14-32

Eckerson W. (2011), Creating an Enterprise Data Strategy. Managing Data as a Corporate Asset. Beye Network. 26-39

Folmer E., Wu H., Van-Bekkum M., (2014), Data standards quality measured for achieving enterprise interoperability: the case of the SETU standard for flexible staffing. Information Systems and e-Business Management. 12(4) 517–541, DOI 10.1007/s10257-014-0236-5

Experian QAS (2008), 10 top Tips for Data Management. Nottingham: QAS Limited http://docs.media.bitpipe.com/io_10x/io_107088/item_597444/10%20Top%20Tips%20for%20Data%20Management.pdf, Last accessed: May 2017

Experian QAS (2008), Transform your data strategy Transform your business results: Guide to laying the foundations for an effective data quality strategy. Experian, Last accessed: April 2017

Experian QAS (2012), Enabling efficiency through Data Governance: a phased approach. Nottingham: QAS Limited, Last accessed: April 2017

Ferguson M. (2007), Accelerating Enterprise Data Governance: Part 1: Foundations. Wilmslow: Intelligent Business Strategies, pp. 37-52

Fisher, T. 2009. The data asset: Govern your data for business success. New York: John Wiley.

Franco J.M. (2014), Master Data Management for Customer Data. Talend, https://www.slideshare.net/jmfranco/mdm-for-customer-data-with-talend , Last accessed: September 2015

Franco J.M. (2014), Master Data Management for Employee Data. Talend, https://www.slideshare.net/jmfranco/mdm-for-employee-data-with-talend, Last accessed: September 2015

Franco J.M. (2014), Master Data Management for ERP Systems. Talend, https://www.slideshare.net/jmfranco/mdm-for-erp-data-with-talend, Last accessed: September 2015

Franco J.M. (2014), Master Data Management for Product Data. Talend, https://www.slideshare.net/jmfranco/mdm-for-product-data-with-talend, Last accessed: September 2015

Griffin J. (2005), Critical Questions for your Master Data Management Initiative. DM Review; New York Vol. 15, Iss 12.

Götz G., Open Source Master Date Management. IAITested http://info.talend.com/rs/talend/images/WP_EN_MDM_IAIT_ProductReview.pdf, Last accessed: May 2016

Götz G., Testing: Talend Data Quality v4. Completely Updated Databases. IAITested http://info.talend.com/rs/talend/images/WP_EN_MDM_IAIT_ProductReview.pdf, Last accessed: May 2016

Guess A.R. (2011), Root causes for Data Quality Problems, http://www.dataversity.net/rootcauses-of-data-quality-problems/, Last accessed: November 2013

Gulipalli G. (2015), Three Major Causes of Enterprise Data Quality Problems, https://www.invensis.net/blog/data-processing/three-major-causes-enterprise-data-quality-problems , Last Accessed: September 2016

Harte-Hanks Trillium Software (2007), Methodology for Enterprise Data Quality and Data Governance. Harte-Hanks Trillium Software. 4-11

Harte-Hanks Trillium Software (2008), Data Quality Essentials: For any Data-Intensive Project. HarteHanks Trillium Software. 9-24

Harte-Hanks Trillium Software (2014), Creating a Single Customer View: The Importance of Data Quality for CRM. Harte-Hanks Trillium Software, https://www.trilliumsoftware.com/crm/white-paper-creating-single-customer-view-importance-data-quality-crm , Last Accessed: September 2016

Harte-Hanks Trillium Software (2014), Evolving from Application to Enterprise Data Quality: A Practical Guide. Harte-Hanks Trillium Software,

Harte-Hanks Trillium Software (2014), Create A Single Customer View. Harte-Hanks Trillium Software,

Harte-Hanks Trillium Software (2014), Creating High-Quality Master Data. Harte-Hanks Trillium Software ,

Hayler A. (2010), Building a Robust Business Case for High Quality Master Data. Intelligent Solutions, Inc.,

Hong D., Zhang Y., Luo J., Liu C., Xu M., Suo Z., (2017), Research on Centralized Data-Sharing Model Based on Master Data Management, MATEC Web of Conferences. DOI: 10.1051/matecconf/201713900195

House of Brick Technologies (2014), Why Data Quality Matters, http://houseofbrick.com/why-dataquality-matters/ , Last accessed: November 2014

Howard P. (2007), Data Migration. Towcester: Bloor Research,

Howard P. (2008), Data Integration Platforms - Talend. London: Bloor Research,

Howard P. (2010), Loading and analysing web data. Considerations and recommendations. London: Bloor Research ,

Howard P. (2010), Talend Data Management Platform. London: Bloor Research,

Howard P. & Stanley N. (2010), Solvency II: data quality and governance. London: Bloor Research,

IDG Research Services (2012), Data Visualization: Making Big Data Approachable and Valuable: A Research Report detailing how Organizations are using data visualization to succeed with big data. SAS,

Imhoff C. (2009), Open Sesame: Why Open Source BI, Data Integration, and Data Warehousing Solutions are Gaining in Acceptance. Intelligent Solutions, Inc.,

Incisive Media (2012), Building a business case for better data management: Identifying the barriers to improved compliance, efficiency and savings. Talend,

Information Builders (2011), Seven Steps to Effective Data Governance. New York: Information Builders,

Info-Tech Research Group (4th March 2014), Storyboard: Develop a Master Data Management Strategy and Roadmap, InfoTech Research Group, https://www.infotech.com/research/storyboarddevelop-a-master-data-management-strategy-and-roadmap , Last accessed: May 2014

Inmon W.H. (2007), The Evolution of Integration. Inmon Consulting Services, Inc, https://www.keyinfo.com/wp-content/uploads/2015/04/The-Evolution-of-Integration-by-W.-H.-Inmon.pdf , Last Accessed: December 2019

Iversen H.K. (2016), Major causes of Enterprise Data Quality Problems, http://www.xangati.com/blog/7-major-causes-of-enterprise-data-quality-problems/ , Last accessed: June 2016

Jones S., Ross S., Ruusalepp R. & Dobreva-McPherson M. (2009) Data Audit Framework Methodology, DO - 10.13140/RG.2.1.3389.0969

Joshi A. (2007), "MDM Governance: A unified team approach." Cutter IT Journal, Vol. 20. 30-35

Kopcke J. (2008) Master Data Management: Old Problem with a new urgency http://www.oracle.com/us/products/middleware/bus-int/064333.pdf , Last accessed: September 2014

Koltay, T. (2016). Data governance, data literacy and the management of data quality. IFLA Journal, 42(4), 303-312., DOI: 10.1177/0340035216672238

Ladley J. (2012). Data Governance. How to design, Deploy, and sustain an effective Data Governance program. MK Publications. pp. 162-165.

Limpert E. & Stahel W. (2011) Problems with Using the Normal Distribution – and Ways to Improve Quality and Efficiency of Data Analysis. PLoS ONE 6(7): e21403. doi:10.1371/journal.pone.0021403

Lo A. (22nd September 2016), Introduction to Azure Data Factory Service, a data integration service in the cloud, Microsoft Co.

Lorh, S. (2011) When there's no such a thing as too much information. New York Times (April 23).

Loshin D (2009), Master Data management space issue. Amsterdam: Elsevier/Morgan Kaufmann.

Loshin D. (2010), Buyer's Guide: Choosing data quality tools and software, https://searchdatamanagement.techtarget.com/news/2240025847/Buyers-Guide-Choosing-data-quality-tools-and-software, Last accessed: September 2014

Loshin D. (2010), Three Fundamental Techniques to Maximize the Value of Your Enterprise Data. Knowledge Integrity, Inc., http://docplayer.net/9678762-Three-fundamental-techniques-to-maximize-the-value-of-your-enterprise-data.html, Last accessed: February 2012

Loshin D. (2011), Busting 7 Myths about Data Quality Management. Knowledge Integrity, Inc., http://docplayer.net/2546070-Busting-7-myths-about-master-data-management.html, Last accessed: February 2012

Loshin D. (2011), Populating a Data Quality Scorecard with Relevant Metrics. DataFlux Corporation, pp. 3-11

Loshin, D. (2011). The practitioners guide to data quality improvement. Burlingtonm MAL Morgan Kaufmann.

Madsen M. R. (2009), The Role of Open Source in Data Integration. Third Nature. 6-16

Mahanti R (2018), Data Quality: dimensions, measurement, strategy, management and goverane, American Society for Quality, Quality Press

Maydanchick A. (2007), Data Quality Assessment, Technics Publications. 39-54, 243-270

McKenna B. (2011), Data quality improvement impeded by lack of automation. https://www.computerweekly.com/news/2240039405/Data-quality-improvement-impeded-by-lack-of-automation, Last accessed: November 2016

McKnight W. (2009), 7 Sources of Poor Data Quality, https://www.melissadata.com/enews/articles/0611/2.htm , Last accessed: September 2016

Feldman M, Even A and Parmet Y. (2018) A methodology for quantifying the effect of missing data on decision quality in classification problems. Commun Stat Theory Methods 47: 2643–2663., DOI: 10.1080/03610926.2016.1277752

Myers M.D. (2008) Qualitative Research in Business & Management, SAGE Publications,

Novak S., (19th July 2016), Initial Data Assessment Activities, https://www.datamigrationfundamentals.com/posts/initial-data-assessment-activities , Last accessed: April 2017

Oberhofer M. & Dreibelbis A. (2008), An Introduction to the Master Data Management Reference Architecture. http://www.ibm.com/developerworks/data/library/techarticle/dm-0804oberhofer/, IBM DeveloperWorks, Last accessed: February 2017

Ohbyung K., Namyeon L., Bongsik S., (2014), Data quality management, data usage experience and acquisition intention of big data analytics, International Journal of Information Management Volume 34, Issue 3, June 2014, Pages 387-394 Science Direct, https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Olofson C. (2008), Talend Uses Open Source to Deliver Low-Cost, Easy-to-Use Enterprise Data Integration. Framingham: IDC

Owens J. (2012), Seven Fatal Errors in MDM and how to avoid them. https://hubdesignsmagazine.com/2012/11/05/seven-fatal-errors-in-mdm-and-how-to-avoid-them-byjohn-owens/ , Last accesed: March 2014

Power D. & Hunt J. (2013), The 8 Worst Practices in Master Data Management and How to Avoid Them. Hingham: Hub Designs

Quinn K. R. (2013), Worst Practices in Business Intelligence Why BI Applications Succeed Where BI Tools Fail. New York: Information Builders

Robert R. (2010), Enterprise Customers Embrace Deduplication Technology. Framingham: IDC Go-to-Market Services

Rowe N. (2012), The State of Master Data Management, 2012: Building the foundations for a better Enterprise. Kennesaw: Stibo Systems, Inc.

Russom P. (2006), Taking Data Quality to the Enterprise through Data Governance. Seattle: TDWI

Russom P. (2010), Introduction to Unified Data Model. Seattle: TDWI

Sarsfield S. (2011), Data Governance Imperative, a business strategy for corporate data. IT governance publishing

Sarsfield S. (2011), Top Ten Root Causes of Data Quality Problems: Part Five, http://datagovernance.blogspot.co.uk/2011/08/top-ten-root-causes-of-data-quality_31.html, Last accessed: March 2014

Sarsfield S. (2011), Top Ten Root Causes of Data Quality Problems: Part Four, http://datagovernance.blogspot.co.uk/2011/08/top-ten-root-causes-of-data-quality_30.html, Last accessed: March 2014

Sarsfield S. (2011), Top Ten Root Causes of Data Quality Problems: Part One, http://datagovernance.blogspot.co.uk/2011/08/top-ten-root-causes-of-data-quality.html, Last accessed: March 2014

Sarsfield S. (2011), Top Ten Root Causes of Data Quality Problems: Part Three, http://datagovernance.blogspot.co.uk/2011/08/top-ten-root-causes-of-data-quality_29.html, Last accessed: March 2014

Sarsfield S. (2011), Top Ten Root Causes of Data Quality Problems: Part Two, http://datagovernance.blogspot.co.uk/2011/08/top-ten-root-causes-of-data-quality_25.html, Last accessed: March 2014

Saunders M., Lewis P., & Thornhill A. (2007), Research Method for Business Students. Prentice Hall, 4th Ed.

Saunders M., Lewis P., & Thornhill A. (2012) “Research Methods for Business Students” 6th edition, Pearson Education Limited

Seiner R. (2014), NON – INVASIVE DATA GOVERNANCE. The path of Least Resistance and Greatest Success, Technics Publications LLC, pp. 12-25, 77-87

Sherman R., (4th August 2011). A must to avoid: Worst practices in enterprise data governance, Athena IT solutions, http://searchdatamanagement.techtarget.com/feature/A-must-to-avoid-Worstpractices-in-enterprise-data-governance , Last accessed: April 2014

Simon P. (2011), Enterprise Data Governance: The Human Element. Cary: DataFlux,

Smith A. (2008), The Role of Politics in Data Governance.Enterprise Information Management Institute. Volume 2 October 2008 Edition. http://www.eiminstitute.org/library/eimi-archives/volume-2-issue-6-september-2008-edition/the-role-of-politics-in-data-governance Last accessed: 14/06/2020

Stylianou A., Jeffries C., Robbins S., (1996), Corporate mergers and the problems of IS integration, Information & Management Volume 31, Issue 4, 15 December 1996, Pages 203-213, Science Direct, https://doi.org/10.1016/S0378-7206(96)01082-8

Talend (2009), Practical Open Source Data Integration. Case Studies & Implementation Examples. Volume 1, 2, 3 & 4. Talend

Talend Technical Note (2008), Integrating SAP data in the information system using open source data integration. Talend

Talend Technical Note (2009), High Performance for Intergrating Massive Data Volumes with Talend Integration Suite MPx. Talend

Talend Technical Note (2009), Leveraging Open Source Data Quality—Practical Examples, Illustrated with Talend Data Quality. Talend

Talend White Paper (2010), Master Reference Data: Extract value from your most common data. Talend

Talend White Paper (2010), Top 10 Reasons Why Data Quality Matters. Talend

Talend White Paper (2012), Building a business case for better data management: Identifying the barriers to improved compliance, efficiency and savings. Talend

Talend White Paper (2012), Data Quality Dashboards in Support of Data Governance. Talend

Talend White Paper (2012), Open Source Master Data Management the Time is Right. Talend

Talend White Paper (2012), The Butterfly Effect on Data Quality: How small data quality issues can lead to big consequences. Talend

The Information Difference company Ltd (2009), Master Management Projects in Paractise: An Information Difference Research Study. The Information Difference company Ltd

Umar A., Karabatis G., Linda N., Horowitz B, & Elmagardmid A. (1999) —Enterprise Data Quality: A Pragmatic Approach‖, Information Systems Frontiers (1)3, pp. 279–301

Vilminko – Heikkinen R. & Pekkola S., (2013) Establishing an Organization's Master Data Management Function: A Stepwise Approach, 2013 46th Hawaii International Conference on System Sciences, pp 4722 – 4728

Vincent M. (2014), Implementing Real-Time Data Quality Management: The Best Solution for Fixing Bad Data. New York: Information Builders,

Vincent M. (2014), Optimizing Data Quality in the Enterprise: How to Tackle Your Bad Information. New York: Information Builders,

Viswanathan K. (2007), A Methodology for Sustainable Success with MDM Initiatives. Tata Consultancy Services Limited, http://www.sourcemediaconferences.com/CDISP07/pdf/Vishwanathan.pdf, Last accessed: March 2014

Weber K., Otto B., & Osterle H. (2009). One size does not fit all—a contingency approach to data governance. ACM J. Data Inform. Quality 1, 1, Article 4 (June 2009), 27 pages. DOI = 10.1145/1515693.1515696. http://doi.acm.org/10.1145/1515693.1515696.

White A. & Radcliffe J. (2008) Mastering Master Data Management, Gartner Research, https://www.gartner.com/doc/670209/overview-mastering-master-data-management, Last accessed: March 2014

McKnight, W. (2006). Justifying and implementing master data management for the enterprise. DM Review, 16(4), 12-14.

Wolter R. & Haselden K., (2006), The What, Why, and How of Master Data Management, Microsoft Corporaton, https://msdn.microsoft.com/en-us/library/bb190163.aspx , Last accessed: March 2014

Woodie A. (2016), Avoid these five big data governance mistakes. https://www.datanami.com/2016/06/22/avoid-five-big-data-governance-mistakes, Datanami, Last accessed: February 2017

Zachman J. (2008), DAMA – GUIDE TO THE DATA MANAGEMENT BODY OF KNOWLEDGE Technics Publications LLC. 37-44, 171-192

Zycus (2011) Item Master Enrichment: A Business Case. Zycus, https://www.scribd.com/document/204575501/Business-Case-Item-Master-Enrichment, Last accessed: March 2014

THANK YOU FOR READING THIS THESIS

Appendix 1.
Interview Questions 23, 25, 26, 27, 28

Please refer to page 5 to page 34 at Volume 2

Appendix 2.
Company One Supporting documents

Please refer to page 35 to page 39 at Volume 2

Appendix 3.
Company Two Supporting documents

Please refer to page 40 to 42 at Volume 2

Appendix 4.
Interview Answers in Gridviews

Please refer to page 43 to 145 at Volume 2

Appendix 5.
Interview Document

Please refer to page 146 to 151 at Volume 2

Appendix 6.
Student Ethics Checklist

Please refer to page 152 to page 153 at Volume 2

Appendix 7.
Practitioner's recommendations

Please refer to page 153 to 155 at Volume 2

Appendix 8
Technical Further Work

Please refer to page 156 to 158

Master Data Management: Its importance and reasons for failed implementations

Panagiotis Lepeniotis

A thesis submitted in partial fulfilment of the requirements of

Sheffield Hallam University

for the degree of Doctor of Philosophy

Volume 2

<u>This document contains only the Appendices</u>

January 2020

Candidate Declaration

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.

2. None of the material contained in the thesis has been used in any other submission for an academic award.

3. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.

4. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.

5. The word count of the thesis is 73,600

Name	*Panagiotis Lepeniotis*
Date	*January 2020*
Award	*PhD*
Faculty	*SHU Faculty of Arts, Computing, Engineering and Sciences*
Director(s) of Studies	*Dr Simon Polovina & Dr Sohrab Moshiri*

Table of Contents

Appendix 1.

Interview Question 23.

Explanation and reasoning of question 23.

Importance of Master Data Management to Core business operations	The interviewer was trying to understand how the interviewees are rating the Master Data Management with regards to their business operations and how the quality of their master data, affect their business process.
Buy-in from Senior Management	For each project and organisation, in general, there needs to be a harmonised cooperation between the IT Department which is usually the department that needs to deliver to Business Requests as well as the Business in general who sets up the Business processes as well as operational processes. The interviewer was trying to understand the views between the three different groups that were taking part in this interview. These three groups were divided in: • IT from a development point of view, Business • Business Process Owners, Senior Management and Decision Makers • Project Management, Solution Architects and Process Optimisers.
Resources/ Budget support for MDM	Even if there was a project of Master Data Management or an element of a Data Quality project task within the larger scale of a business Transformation process and even if there was support for this streamline, the interviewer wanted to understand how the business and the decision-makers were supporting the process of Data Quality insurance and Master Data Management in practice. Bearing in mind that in business everything counts only when monetary values are assigned next to each task or process. That is why the interviewer wanted to understand what is the level of support by the business on these project tasks as well as what was the general feeling that the interviewees were passing to the interviewer.
Adherence to Master Data policies	The interviewer with this measure wanted to understand the level of engagement from both business or general employees that are responsible for executing the business process as well as the IT

	people. The interviewer, during the conversation, wanted to understand the processes of engaging in this polices as a project, as a system and as the process. In case of Business Transformation project, the interviewer wanted to analyse the “AS-IS” process and the “TO BE” processes and how both sides of the project are adapting on the Master Data Policy.
Trust in Master Data currently	Even if it was a Business Transformation Project or a current Data Quality / Data Governance policy that it was already in place, the interviewer wanted to understand what is the level of Trust on the current Master Data as well as what would be the trust after a policy exclusively developed for Data Management including both Static and Transactional data from both the interviewees but more specifically the Senior Management and the decision-makers. The interviewer wanted to understand that even if there was an investment from the business at some point regarding the Data Quality Management was giving peace of mind to the processes or if there was going to increase the level of trust on the processes after implementation of a project or policy of this nature.
Trust in Data systems and Policies currently	Following the above, except for the actual data that is involved in any bidirectional integration defining a business process. It is the system that gives the user the tool to apply and follow this policy as well as to commit any transactions driven by the policy. That is why the interviewer wanted to understand the challenges that the organisations were facing with regards to systems. This measure was giving two different metrics. If the mark was high, the interviewer wanted to understand how the trust has been achieved and if the mark was low, the discussion would focus on the reasons why the level of trust is low.

How likely is to identify Records with significant errors currently	Based on the mark of this question given by the interviewee, the interviewer wanted to drill down even further. First of all, to understand if it is easy or difficult to measure the quality of the data. How the people were being able to separate the correct from the errored record. And what was the business process which defines the identification of the right and wrong as well as at what level the people were in the position to understand any discrepancies and what was the impact on both right and wrong cases? One of the important parts that the interviewer wanted to understand as well as the impact that it had been caused (If any) based on the time that it took to the business to realise the data error. How much would differ if any error would have been identified earlier?
How likely is to fix these errors in less than an hour currently	With this measure question, the interviewer wanted to understand in more detail not only the actual time that it takes the users to fix an error but also in conjunction with the previous metric to understand what is the impact of late resolution to a potential data quality issue. In the type of industries that are depended on data-heavy processes what is the usually required service level agreement with customers either internal or external and what happens when this resolution will not occur within the agreed or required time. During the conversation, one of the other metrics that are not officially recorded was the frequency of these events and how the error resolution had evolved during once a discrepancy would have been identified.
How likely is for the information to be gained faster currently	This question was focused more on the Data Management side of the process than the service level. So, the main conversation was around how the correct policy, as well as systems in addition to a proper governance on any integration between systems, was set up on strong foundations regarding the data management, was helping in getting the information out of the data Faster in opposition to an information-seeking process that doesn't have any of the above. Of course, the trust in the data, as well as the systems in existing programmes or business transformation programmes, should have been identified at the beginning or in what level this process has been set up.

How likely is a decision to be made faster currently	Similarly, with the above question, the interviewer wanted to understand if the information was being used to make any business decisions based on the data. This decision was again depending on the level of Trust that the Data Governance was followed as well as the level of trust on the systems that produce the data, as a result, the information and of course that the systems comply with the Governance and Data Policies that had been set up. If there were any policies.
How likely is to integrate new data sources faster currently	This question was mainly focus on the architecture of the Data Management Process. Based on the metric that the interviewees would be given, the interviewer would begin a debate on how the organisation or the project or the process was bringing together information from multiple systems if they had specific rules and specific Data Dictionaries that they should follow and how they were making use of these processes in an existing Data Ecosystem or if they were taking into consideration any reusable processes for any future integrational environment.
How likely is for the employee to reduce	The last metric value of the question twenty-three was mainly industry and department related. The interviewer had to
the time that he spends on data-centric processes currently	understand from the conversation how the data are captured by a manual input process and how the user gets the answers to the data related questions that interrogate the systems or the tools that access the data. How the quality of data affects the time that the user spends and how accurate is at the end the process that the user commits.

Table 1. Question 23

Question	Role	Function	Answer
Importance of Master Data Management to Core business operations	CIO	IT	• CIO1: **10** • CIO2: **10**
	CFO	Business	• CFO1: **6** • CFO2: **5** • CFO3: **8**
	MDM Directors	IT/Business	• MDM Director: **10**
	Head of Development	IT	• Head of Dev 1: **8** • Head of Dev 2: **9**
	Business Transformation Program Director	Business	• BTPD 1: **6** • BTPD 2: **4** • BTPD 3: **5**
	IT Project Manager	IT	• IT PM 1: **7**

			• IT PM 2: **7**
	Head of Business Analysis	IT/Business	• Head of BA: **6**
	Business Analysts	IT/Business	• BA 1: **6** • BA 2: **4** • BA 3: **4** • BA 4: **5**
Buy-in from Senior Management	CIO	IT	• CIO1: **10** • CIO1: **10**
	CFO	Business	• CFO1: **8** • CFO2: **8** • CFO3: **7**
	MDM Directors	IT/Business	• MDM Director: **10**
	Head of Development	IT	• Head of Dev 1: **8** • Head of Dev 2: **9**
	Business Transformation Program Director	Business	• BTPD 1: **7** • BTPD 2: **8** • BTPD 3: **4**
	IT Project Manager	IT	• IT PM 1: **8** • IT PM 2: **7**
	Head of Business Analysis	IT/Business	• Head of BA: **6**
	Business Analysts	IT/Business	• BA 1: **9** • BA 2: **7** • BA 3: **7** • BA 4: **6**
Resources/ Budget support for MDM	CIO	IT	• CIO1: **10** • CIO2: **10**
	CFO	Business	• CFO1: **9** • CFO2: **8** • CFO3: **8**
	MDM Directors	IT/Business	• MDM Director: **10**
	Head of Development	IT	• Head of Dev 1: **9** • Head of Dev 2: **9**
	Business Transformation Program Director	Business	• BTPD 1: **6** • BTPD 2: **4** • BTPD 3: **5**
	IT Project Manager	IT	• IT PM 1: **8** • IT PM 2: **8**
	Head of Business Analysis	IT/Business	• Head of BA: **8**
	Business Analysts	IT/Business	• BA 1: **8** • BA 2: **7** • BA 3: **7** • BA 4: **6**
Adherence to Master Data policies	CIO	IT	• CIO1: **10** • CIO1: **10**
	CFO	Business	• CFO1: **9** • CFO2: **9** • CFO3: **8**

	MDM Directors	IT/Business	• MDM Director: **10**
	Head of Development	IT	• Head of Dev 1: **9** • Head of Dev 2: **9**
	Business Transformation Program Director	Business	• BTPD 1: **8** • BTPD 2: **8** • BTPD 3: **6**
	IT Project Manager	IT	• IT PM 1: **9** • IT PM 2: **8**
	Head of Business Analysis	IT/Business	• Head of BA: **8**
	Business Analysts	IT/Business	• BA 1: **8** • BA 2: **8** • BA 3: **7** • BA 4: **7**
Trust in Master Data currently	CIO	IT	• CIO1: **10** • CIO1: **10**
	CFO	Business	• CFO1: **9** • CFO2: **9** • CFO3: **8**
	MDM Directors	IT/Business	• MDM Director: **10**
	Head of Development	IT	• Head of Dev 1: **10** • Head of Dev 2: **10**
	Business Transformation Program Director	Business	• BTPD 1: **9** • BTPD 2: **9** • BTPD 3: **7**
	IT Project Manager	IT	• IT PM 1: **8** • IT PM 2: **8**
	Head of Business Analysis	IT/Business	• Head of BA: **7**
	Business Analysts	IT/Business	• BA 1: **8** • BA 2: **7** • BA 3: **8** • BA 4: **6**
Trust in Data systems and Policies currently	CIO	IT	• CIO1: **7** • CIO1: **7**
	CFO	Business	• CFO1: **6** • CFO2: **6** • CFO3: **5**
	MDM Directors	IT/Business	• MDM Director: **6**
	Head of Development	IT	• Head of Dev 1: **6** • Head of Dev 2: **7**
	Business Transformation Program Director	Business	• BTPD 1: **5** • BTPD 2: **6** • BTPD 3: **8**
	IT Project Manager	IT	• IT PM 1: **6** • IT PM 2: **5**
	Head of Business Analysis	IT/Business	• Head of BA: **5**
	Business Analysts	IT/Business	• BA 1: **5**

			• BA 2: **6** • BA 3: **5** • BA 4: **4**
How likely is to identify Records with significant errors currently	CIO	IT	• CIO1: **6** • CIO2: **6**
	CFO	Business	• CFO1: **4** • CFO2: **4** • CFO3: **3**
	MDM Directors	IT/Business	• MDM director: **4**
	Head of Development	IT	• Head of Development 1: **4** • Head of Development 2: **5**
	Business Transformation Program Director	Business	• BTPD 1: **5** • BTPD 2: **4** • BTPD 3: **6**
	IT Project Manager	IT	• IT PM 1: **5** • IT PM 2: **4**
	Head of Business Analysis	IT/Business	• Head of Business Analysis: **4**
	Business Analysts	IT/Business	• BA 1: **3** • BA 2: **4** • BA 3: **4** • BA 4: **4**
How likely is to fix these errors in less than an hour currently	CIO	IT	• CIO1: **3** • CIO2: **3**
	CFO	Business	• CFO1: **3** • CFO2: **3** • CFO3: **3**
	MDM Directors	IT/Business	• MDM director: **4**
	Head of Development	IT	• Head of Development 1: **4** • Head of Development 2: **4**
	Business Transformation Program Director	Business	• BTPD 1: **3** • BTPD 2: **4** • BTPD 3: **7**
	IT Project Manager	IT	• IT PM 1: **4** • IT PM 2: **4**
	Head of Business Analysis	IT/Business	• Head of Business Analysis: **4**
	Business Analysts	IT/Business	• BA 1: **4** • BA 2: **4** • BA 3: **3** • BA 4: **4**
How likely is for the information to be gained faster currently	CIO	IT	• CIO1: **4** • CIO1: **4**
	CFO	Business	• CFO1: **4** • CFO2: **4** • CFO3: **5**
	MDM Directors	IT/Business	• MDM Director: **5**
	Head of Development	IT	• Head of Dev 1: **5**

			• Head of Dev 2: **5**
	Business Transformation Program Director	Business	• BTPD 1: **4** • BTPD 2: **5** • BTPD 3: **7**
	IT Project Manager	IT	• IT PM 1: **4** • IT PM 2: **4**
	Head of Business Analysis	IT/Business	• Head of BA: **5**
	Business Analysts	IT/Business	• BA 1: **4** • BA 2: **5** • BA 3: **4** • BA 4: **4**
How likely is a decision to be made faster currently	CIO	IT	• CIO1: **3** • CIO2: **3**
	CFO	Business	• CFO1: **6** • CFO2: **6** • CFO3: **8**
	MDM Directors	IT/Business	• MDM Director: **3**
	Head of Development	IT	• Head of Dev 1: **3** • Head of Dev 2: **4**
	Business Transformation Program Director	Business	• BTPD 1: **5** • BTPD 2: **5** • BTPD 3: **6**
	IT Project Manager	IT	• IT PM 1: **2** • IT PM 2: **3**
	Head of Business Analysis	IT/Business	• Head of BA: **4**
	Business Analysts	IT/Business	• BA 1: **4** • BA 2: **3** • BA 3: **4** • BA 4: **3**
How likely is to integrate new data sources faster currently	CIO	IT	• CIO1: **6** • CIO2: **7**
	CFO	Business	• CFO1: **8** • CFO2: **8** • CFO3: **7**
	MDM Directors	IT/Business	• MDM Director: **6**
	Head of Development	IT	• Head of Dev 1: **6** • Head of Dev 2: **7**
	Business Transformation Program Director	Business	• BTPD 1: **4** • BTPD 2: **5** • BTPD 3: **7**
	IT Project Manager	IT	• IT PM 1: **4** • IT PM 2: **4**
	Head of Business Analysis	IT/Business	• Head of BA: **7**
	Business Analysts	IT/Business	• BA 1: **6** • BA 2: **5** • BA 3: **4** • BA 4: **6**

How likely is for the employee to reduce the time that he spends on data-centric processes currently	CIO	IT	• CIO1: **4** • CIO2: **4**
	CFO	Business	• CFO1: **4** • CFO2: **5** • CFO3: **4**
	MDM Directors	IT/Business	• MDM Director: **5**
	Head of Development	IT	• Head of Dev 1: **5** • Head of Dev 2: **5**
	Business Transformation Program Director	Business	• BTPD 1: **5** • BTPD 2: **5** • BTPD 3: **7**
	IT Project Manager	IT	• IT PM 1: **5** • IT PM 2: **5**
	Head of Business Analysis	IT/Business	• Head of BA: **5**
	Business Analysts	IT/Business	• BA 1: **6** • BA 2: **5** • BA 3: **5** • BA 4: **4**

Interview Question 25.

Explanation and Reasoning of question 25

Process	Standardized training for master data system	With this metric, the interviewer wanted to understand how the training for the process is delivered, if there was any training and how effective this training was. The key element for the training that the interviewer wanted to absorb from this metric was basically if there was a specific method of delivering this training. In details, if there was a specific training session for specific audiences if the objectives were set before the training again, based on the audience. If there was a different approach with regards the part of the system, the users for each part and if there were different sections assigned to specific users that would make the training more effective.

	End-user needs for data access and use collected	Following the above metric, the interviewer wanted to understand if before any training was a structured requirement analysis based on the roles of the users. The background knowledge of the users that would take the responsibility of receiving the training but more importantly managing the process in a business as usual basis. Also, one of the important parts that the interviewer wanted to understand was the fact that the roles and responsibilities of each user group that they would use or already use a master data system or steward a data governance policy on an existing system could fully understand their responsibilities, their boundaries and their influence within the organisation.
Organization	Executive sponsor for MDM	With this metric, the interviewer wanted to understand from the interviewees, first of all, if there was an executive sponsor for the Master Data Management Project or Data Quality process of Data Governance policy on the existing systems, a new project or the Business Transformation programme. And of course, the extra important part that the interviewer wanted to understand as if the level of support that the project team or the Data The management team was receiving from the executive sponsor was the desired one. Also, one other important element that the
		the interviewer wanted to understand during the conversation was the level of alignment between the Executive sponsor and the Senior Management/ board of directors and also the level in the hierarchy within the business that the executive sponsor was in.
	Cross-functional MDM team	The main outcome that the interviewer wanted to get out of this question was the roles between the Master Data Management Project and how effective and clear roles are distributed between the team. Also, the experience before each person that now would belong to a Master Data Management team should be evaluated and a “fit for purpose” approach would be assessed based on the level of knowledge of each person that form this team.

	Defined Create, Read, Update and Delete (CRUD) role	Of course, Training should be given to all the member of the organisation but basically around the Data Governance and how each department should respect the data that they capture, update, delete or read. The interviewer wanted to capture with this question the level of understanding across all the departments of the organisation how each function uses the data and if they are educated enough with regards to the policy but most importantly if these functions within the organisation could understand the value and the importance of this policy. And if these departments were aware of how much their Business as usual job is depended on this policy.
Knowledge	Discovery and identification of all business data	With this question, the interviewer wanted to understand if the interviewees were aware first of all with the term of Business Data, and operational knowledge. Since the interviewees were covering all the levels of an organisation, the interviewer wanted to understand the general understanding of the business processes which this translates to business knowledge which results in business information. The interviewer wanted to understand the level of knowledge that these people were holding and how this knowledge that they were holding was matching the process that they had to deliver on their project, their business transformation programme or their Business as usual everyday tasks.
	Classification and definition of all business Data	The main outcome of this metric is highly connected with the metric just above. Following the previous question, the interviewer wants to understand if the users understand the different classifications between each function of the business but also within the same function. By Classification of Business Data, the interviewer is referring to a specific set of Static data that are used for a specific business process. Following that, the important
		the metric that the interviewer wants to get out of this question is the level of understanding of this set of data as well as the level of understanding that involved in the management of this set of data.

Performance	Measurement tools to track and report data quality	The key metric that the interviewer wants to take of this question is the understanding by the interviewees of how important is the quality of the data that they have to operate with on every function of the organisation and every level of their project or programme. By the word “understanding”, the interviewer is referring on the way that the interviewees measure the quality of the data that they have to work within a daily basis as well as the dataset that they will have to take on, on any new project or programme. That is why the interviewer tries with this question to understand how the quality of each data is measured and the results of this measure are used.
	ROI for MDM defined and Tracked	With this question, the interviewer tries to understand from the interviewees as long as they have a Master Data Management system, process or Data Governance policy in place how they measure the Return on the investment that the organisation has done on having a system or process like that. Or if the organisation that the interviewee is coming from not having a Master Data Management in place, what is the Return on their investment that they are expecting once they have a system like that or a data governance policy in place.
	End-Users time to access master data tracked and measured	With this metric, the interviewer wanted to understand how the interviewees see the benefits (if any) after implementing a Master Data Management solution from an end-user perspective. The interviewer asked the interviewees to estimate the time that it takes to the business as usual employees to process the data that they need with the current solution that is in place, the estimation of the time after a solution has been implemented and if the trust on the data that they use reduces the time to complete a process or not.

Table 2 Question 25

Role	Type	Details	Answer
CIO	Process	Standardized training for master data system	• CIO1: **3** • CIO2: **4**
		End-user needs for data access and use collected	• CIO1: **4** • CIO2: **4**
	Organization	Executive sponsor for MDM	• CIO1: **3** • CIO2: **4**
		Cross-functional MDM team	• CIO1: **3** • CIO2: **3**

		Defined Create, Read, Update and Delete (CRUD) role	• CIO1: **3** • CIO2: **2**
	Knowledge	Discovery and identification of all business data	• CIO1: **2** • CIO2: **2**
		Classification and definition of all business Data	• CIO1: **2** • CIO2: **3**
	Performance	Measurement tools to track and report data quality	• CIO1: **4** • CIO2: **4**
		ROI for MDM defined and Tracked	• CIO1: **1** • CIO2: **3**
		End-Users time to access master data tracked and measured	• CIO1: **1** • CIO2: **3**
CFO	Process	Standardized training for master data system	• CFO1: **3** • CFO2: **2** • CFO3: **3**
		End-user needs for data access and use collected	• CFO1: **3** • CFO2: **3** • CFO3: **4**
	Organization	Executive sponsor for MDM	• CFO1: **3** • CFO2: **5** • CFO3: **4**
		Cross-functional MDM team	• CFO1: **5** • CFO2: **5** • CFO3: **4**
		Defined Create, Read, Update and Delete (CRUD) role	• CFO1: **3** • CFO2: **5** • CFO3: **5**
	Knowledge	Discovery and identification of all business data	• CFO1: **1** • CFO2: **4** • CFO3: **3**
		Classification and definition of all business Data	• CFO1: **2** • CFO2: **2** • CFO3: **1**
	Performance	Measurement tools to track and report data quality	• CFO1: **2** • CFO2: **2** • CFO3: **1**
		ROI for MDM defined and Tracked	• CFO1: **2** • CFO2: **1** • CFO3: **1**
		End-Users time to access master data tracked and measured	• CFO1: **1** • CFO2: **2** • CFO3: **2**
MDM Director	Process	Standardized training for master data system	• MDM Director: **7**
		End-user needs for data access and use collected	• MDM Director: **5**
	Organization	Executive sponsor for MDM	• MDM Director: **5**
		The cross-functional MDM team	• MDM Director: **6**
		Defined Create, Read, Update and Delete (CRUD) role	• MDM Director: **5**
	Knowledge	Discovery and identification of all business data	• MDM Director: **5**
		Classification and definition of all business Data	• MDM Director: **6**

	Performance	Measurement tools to track and report data quality	• MDM Director: **3**
		ROI for MDM defined and Tracked	• MDM Director: **4**
		End-Users time to access master data tracked and measured	• MDM Director: **4**
Head of Development	Process	Standardized training for master data system	• Head of Dev 1: **3** • Head of Dev 2: **2**
		End-user needs for data access and use collected	• Head of Dev 1: **2** • Head of Dev 2: **2**
	Organization	Executive sponsor for MDM	• Head of Dev 1: **2** • Head of Dev 2: **3**
		The cross-functional MDM team	• Head of Dev 1: **2** • Head of Dev 2: **1**
		Defined Create, Read, Update and Delete (CRUD) role	• Head of Dev 1: **1** • Head of Dev 2: **1**
	Knowledge	Discovery and identification of all business data	• Head of Dev 1: **1** • Head of Dev 2: **2**
		Classification and definition of all business Data	• Head of Dev 1: **2** • Head of Dev 2: **2**
	Performance	Measurement tools to track and report data quality	• Head of Dev 1: **2** • Head of Dev 2: **2**
		ROI for MDM defined and Tracked	• Head of Dev 1: **1** • Head of Dev 2: **1**
		End-Users time to access master data tracked and measured	• Head of Dev 1: **1** • Head of Dev 2: **1**
Business Transformation Program Director	Process	Standardized training for master data system	• BTPD 1: **3** • BTPD 2: **2** • BTPD 3: **4**
		End-user needs for data access and use collected	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **5**
	Organization	Executive sponsor for MDM	• BTPD 1: **3** • BTPD 2: **3** • BTPD 3: **4**
		Cross-functional MDM team	• BTPD 1: **2** • BTPD 2: **3** • BTPD 3: **5**
		Defined Create, Read, Update and Delete (CRUD) role	• BTPD 1: **5** • BTPD 2: **5** • BTPD 3: **7**
	Knowledge	Discovery and identification of all business data	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **5**
		Classification and definition of all business Data	• BTPD 1: **4** • BTPD 2: **3** • BTPD 3: **5**
	Performance	Measurement tools to track and report data quality	• BTPD 1: **2** • BTPD 2: **3** • BTPD 3: **5**
		ROI for MDM defined and Tracked	• BTPD 1: **2** • BTPD 2: **4**

			• BTPD 3: **5**
		End-Users time to access master data tracked and measured	• BTPD 1: **2** • BTPD 2: **3** • BTPD 3: **5**
IT Project Manager	Process	Standardized training for master data system	• IT PM 1: **3** • IT PM 2: **3**
		End-user needs for data access and use collected	• IT PM 1: **3** • IT PM 2: **3**
	Organization	Executive sponsor for MDM	• IT PM 1: **4** • IT PM 2: **5**
		The cross-functional MDM team	• IT PM 1: **2** • IT PM 2: **3**
		Defined Create, Read, Update and Delete (CRUD) role	• IT PM 1: **2** • IT PM 2: **2**
	Knowledge	Discovery and identification of all business data	• IT PM 1: **2** • IT PM 2: **1**
		Classification and definition of all business Data	• IT PM 1: **1** • IT PM 2: **2**
	Performance	Measurement tools to track and report data quality	• IT PM 1: **1** • IT PM 2: **2**
		ROI for MDM defined and Tracked	• IT PM 1: **2** • IT PM 2: **3**
		End-Users time to access master data tracked and measured	• IT PM 1: **1** • IT PM 2: **1**
Head of Business Analysis	Process	Standardized training for master data system	• Head of BA: **2**
		End-user needs for data access and use collected	• Head of BA: **3**
	Organization	Executive sponsor for MDM	• Head of BA: **3**
		The cross-functional MDM team	• Head of BA: **4**
		Defined Create, Read, Update and Delete (CRUD) role	• Head of BA: **2**
	Knowledge	Discovery and identification of all business data	• Head of BA: **3**
		Classification and definition of all business Data	• Head of BA: **3**
	Performance	Measurement tools to track and report data quality	• Head of BA: **3**
		ROI for MDM defined and Tracked	• Head of BA: **3**
		End-Users time to access master data tracked and measured	• Head of BA: **2**
Business Analysts	Process	Standardized training for master data system	• BA 1: **2** • BA 2: **3** • BA 3: **3** • BA 4: **4**
		End-user needs for data access and use collected	• BA 1: **2** • BA 2: **3** • BA 3: **3** • BA 4: **3**
	Organization	Executive sponsor for MDM	• BA 1: **4** • BA 2: **3** • BA 3: **3** • BA 4: **4**
		The cross-functional MDM team	• BA 1: **3**

			• BA 2: **3** • BA 3: **5** • BA 4: **4**
		Defined Create, Read, Update and Delete (CRUD) role	• BA 1: **3** • BA 2: **4** • BA 3: **4** • BA 4: **4**
	Knowledge	Discovery and identification of all business data	• BA 1: **3** • BA 2: **2** • BA 3: **2** • BA 4: **2**
		Classification and definition of all business Data	• BA 1: **3** • BA 2: **2** • BA 3: **2** • BA 4: **2**
	Performance	Measurement tools to track and report data quality	• BA 1: **3** • BA 2: **2** • BA 3: **3** • BA 4: **2**
		ROI for MDM defined and Tracked	• BA 1: **2** • BA 2: **5** • BA 3: **5** • BA 4: **4**
		End-Users time to access master data tracked and measured	• BA 1: **2** • BA 2: **2** • BA 3: **2** • BA 4: **2**

Interview Question 26.

Explanation and reasoning of question 26.

Baseline	Organisational Buy-in (Senior Management support, Champion of MDM et)	The first part which defines the foundations of any Master Data Management Project or Data Governance Policy is the support of the Organisation and especially the senior management or the board which is the part of the business that will assign a budget to it. What the interviewer wanted to understand with this metric was basically how the business was supporting the project or programme at the beginning as well as the level of commitment that the users were feeling that the business was prepared to assure.
	Understanding the existing Data environment	Before a project or a programme of Master Data Management or Data Governance the policy is about to start, an initial analysis of the current environment needs to take place. What the interviewer wanted to get with this metric was the actual feeling from the interviewees regarding the data landscape within the organisation was understood from all the departments that would have a significant role during the implementation of this process. The interviewer wanted to understand the level of confidence that the interviewees, as well as the business, have on their current data environment as well as how this environment is going to change following the implementation of the new Master Data Management process.
Emerging	Assessing the need for end-users that interacts with MDM	With this metric, the interviewer wanted to measure the level of understanding that interviewees had with regards to the usage of Master Data Management from an end-user perspective. The interviewer was trying to comprehend from the interviewees if the end-users were aware of the fact that the Master Data Management as a process,

		interacts with their everyday tasks. The interviewer wanted to understand if the end-users are realising that even for example when the end-user was performing a business process, the fundamental data they were using to perform this task was based on Master data. How this data was affecting their job and how this data was important for completing this task with success.
	Measuring Data Quality	Following the above metric, the interviewer wanted to understand how the performance of these daily tasks was affected by the quality of this data. Another part of this metric that the interviewer wanted to understand as if the users were aware of what each selection of the data that they were using was meaning. Even if the level of the data quality was acceptable and within the requirements for performing a task, the end-user knowledge of what this data means could affect any form of measuring the Data Quality.
Early adoptions	Measuring time to Access Information	With this metric, the interviewer wanted to understand if any metrics have been taken by the project team or end-users before the implementation of any Master Data Management processor Business Transformation programme which involved Data Quality tasks. The interviewer wanted to understand the effort that has been taken to measure the time and the cost that the end-user needed to complete a task before the implementation and what were the predictions for the same business process tasks or equivalent business process after the implementation.

	Measuring the ROI of MDM	Following the above metric, the interviewer wanted to understand how the recordings that were mentioned on the previous tasks were translated into Return on the investment that the business is planning to
		make or how the process previous to the implementation was measured with regards to the costs to the business.
BIC Differentiations	Standardised Training	The main metric that the interviewer wanted to take out of this question was mainly to understand if the interviewees are aware of the appropriate type of training and how this training planning can be achieved. The interviewer wanted to understand from the score that the interviewees would give to this metric the level of thought that has been put into the various aspects of the Master Data Management process, the alignment with the principles of that the Data Governance Policy indicates as well as the understanding of the business process by the end-users that process the data by completing their everyday tasks.
	Enforcing CRUD (not define but enforcing)	With this question the interviewer wanted to understand how the Data Governance was being applied or strategically defined on how it will be applied when the Master Data Management and Data Governance Policy Program will be live. The interviewer wanted to understand by the mark that the interviewees would give to that question if there were specific rules following any new creation, alteration or deletion of Data if there was a specific audit trail that was capturing the activity of any alteration on the Master data or any prevention of altering this data based on the rules that had been defined by the Data Governance Policy.

Interview Question 27.

Explanation and Reasoning of question 27

Automation	Auto Internal Capture	A score was required by the interviewees to describe the level of automation that was taking place during the process of creation, alteration and deletion of data under the Master Data Management underlining process. The Internal Capture is referring on receiving data from sources that are not managed by the business but the data that these sources generally are used for performing actions of Create, Read, Update and Delete automatically from system to system. An example of this process is when an Electronic Data Interchange (EDI) is taking place between two organisations.
	Auto External Capture	A score was required by the interviewees to describe the level of automation that was taking place from an end-user perspective. The main purpose the question was to understand the level that a system reacts automatically when an end-user tries to Create, Read, Update or Delete any data that are protected by the Master Data Management platform. In more details, the question wants to capture the way that a system which is part of the Master Data Management Ecosystem automatically reacts when an end-user tries to manually update a record which is part of the Master Data Management Monitoring process.

	Auto Indexing - Sorting	One aspect of a Master Data Management Solution is the prevention and resolution of duplicate records by identifying the potential candidate record that is clustered as a duplicate. As part of the automation of the Master Data Management process, is the automatic identification of this candidates record and based on a specific set of process to index them in relationship with the golden records which is the master record.
	Auto Cleansing	Following the description of the above process, the next metric on the automation of a Master Data Management solution is the auto cleansing of duplicated records or candidates. As part of the automation, the interviewer was trying to measure the score that the interviewees would give with regards to auto resolving process of the potential duplicates. The main outcome that the interviewer wanted to get out of this question was how automated is the process of auto resolving duplicated candidates.

Data Management	Data Enrichment	With this metric, the interviewer wanted to understand how the interviewees understand the data enrichment process as part of a Master Data Management and Data Governance Program. Data Enrichment is focusing on the management of the master record and how the process defines the Master Data as complete. Which attributes describe a complete record and what level of completed attributes are required for classifying a golden record as complete.
	Data Governance	Following the above metric. The data governance is the part of a Master Data Management implementation that the policy is applied. With this question, the interviewer wanted to understand the level of governance that is enforced on the enrichment of the golden record.
	Data Deduplication	The data deduplication metric mainly follows the same logic that the question has been described on the automation part of the question twenty-seven but including any manual intervention from the Master Data Management tool users or Data stewards or Data Custodians.

	Data Cleansing	The Data Cleansing metric mainly follows the same logic that the question has been described on the automation part of the question twenty-seven but including any manual action as well. In this case, the cleansing activity will include and mainly be happening by the Master Data Management end-users, Data Stewarts or data Custodians.
	Data Normalisation	With this question, the interviewer wanted to understand how the interviewees were defining the attributes that all together form the golden record. At that stage, the interviewer having a conversation with the interviewee as well as trying to mainly understand the process of how a set of attributes is defined and clustered as the golden record.
Data Access	Data Access Tools	With this question, the interviewer wanted to understand the way that the users access the Master Data Management platform. How easy is or is going to be when the platform is ready to be used and how well do the interviewees think that the tool will work with the automation.
	BI for MDM	With this question, the interviewer wanted to understand the level of insights that the Master Data Management process is going to give. What kind of metrics with regards to Data Quality the users will want to see and if these metrics are part of measuring the Return on the investment that they are planning on doing with a programme like that.

	Internal Collaboration Tools	The interviewer wanted to understand if there is any plan for any collaboration tools that will help with data access from the Master Data Management developed by the IT team of the organisation and why these tools would be developed outside the Master Data Management Platform.
	External collaboration Tools	Same as the metric above, the interviewer wanted to understand if any tools are going to be bought in addition to the Master Data Management Platform and if yes, why this wasn't a part of the initial Master Data Management programme.
	Mobile Access	With this question, the interviewer wanted to understand if there was a desire by the interviewees to have access to the Master Data Management platform through a mobile device and if yes what would be the reason for that and how that could be useful.

Table 3 Question 27

Role	Action Category	Action Type	Current Use	Future Use
CIO	Automation	Auto Internal Capture	• CIO1: **1** • CIO2: **1**	• CIO1: **8** • CIO2: **9**
		Auto External Capture	• CIO1: **1** • CIO2: **1**	• CIO1: **5** • CIO2: **6**
		Auto Indexing - Sorting	• CIO1: **1** • CIO2: **1**	• CIO1: **4** • CIO2: **5**
		Auto Cleansing	• CIO1: **1** • CIO2: **1**	• CIO1: **6** • CIO2: **4**
	Data Management	Data Enrichment	• CIO1: **1** • CIO2: **1**	• CIO1: **6** • CIO2: **7**
		Data Governance	• CIO1: **1** • CIO2: **1**	• CIO1: **8** • CIO2: **7**
		Data Deduplication	• CIO1: **1** • CIO2: **1**	• CIO1: **9** • CIO2: **9**

		Data Cleansing	• CIO1: **1** • CIO2: **1**	• CIO1: **9** • CIO2: **9**
		Data Normalisation	• CIO1: **1** • CIO2: **1**	• CIO1: **7** • CIO2: **8**
	Data Access	Data Access Tools	• CIO1: **1** • CIO2: **1**	• CIO1: **6** • CIO2: **7**
		BI for MDM	• CIO1: **1** • CIO2: **1**	• CIO1: **4** • CIO2: **6**
		Internal Collaboration Tools	• CIO1: **1** • CIO2: **1**	• CIO1: **5** • CIO2: **6**
		External collaboration Tools	• CIO1: **1** • CIO2: **1**	• CIO1: **5** • CIO2: **5**
		Mobile Access	• CIO1: **1** • CIO2: **1**	• CIO1: **4** • CIO2: **4**
CFO	Automation	Auto Internal Capture	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **7** • CFO2: **6** • CFO3: **7**
		Auto External Capture	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **5** • CFO2: **5** • CFO3: **6**
		Auto Indexing - Sorting	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **3** • CFO2: **4** • CFO3: **4**
		Auto Cleansing	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **9** • CFO2: **9** • CFO3: **9**
	Data Management	Data Enrichment	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **8** • CFO2: **8** • CFO3: **9**
		Data Governance	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **9** • CFO2: **9** • CFO3: **9**
		Data Deduplication	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **10** • CFO2: **9** • CFO3: **10**
		Data Cleansing	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **10** • CFO2: **10** • CFO3: **10**
		Data Normalisation	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **7** • CFO2: **7** • CFO3: **6**
	Data Access	Data Access Tools	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **5** • CFO2: **6** • CFO3: **5**
		BI for MDM	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **6** • CFO2: **6** • CFO3: **8**
		Internal Collaboration Tools	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **5** • CFO2: **6** • CFO3: **6**
		External collaboration Tools	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **5** • CFO2: **6** • CFO3: **6**

		Mobile Access	• CFO1: **1** • CFO2: **1** • CFO3: **1**	• CFO1: **4** • CFO2: **4** • CFO3: **6**
MDM Director	Automation	Auto Internal Capture	• MDM Director: **1**	• MDM Director: **10**
		Auto External Capture	• MDM Director: **1**	• MDM Director: **9**
		Auto Indexing - Sorting	• MDM Director: **1**	• MDM Director: **9**
		Auto Cleansing	• MDM Director: **1**	• MDM Director: **10**
	Data Management	Data Enrichment	• MDM Director: **1**	• MDM Director: **8**
		Data Governance	• MDM Director: **1**	• MDM Director: **10**
		Data Deduplication	• MDM Director: **1**	• MDM Director: **10**
		Data Cleansing	• MDM Director: **1**	• MDM Director: **10**
		Data Normalisation	• MDM Director: **1**	• MDM Director: **9**
	Data Access	Data Access Tools	• MDM Director: **1**	• MDM Director: **9**
		BI for MDM	• MDM Director: **1**	• MDM Director: **10**
		Internal Collaboration Tools	• MDM Director: **1**	• MDM Director: **9**
		External collaboration Tools	• MDM Director: **1**	• MDM Director: **9**
		Mobile Access	• MDM Director: **1**	• MDM director: **5**
Heads of Development	Automation	Auto Internal Capture	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev1: **9** • Head of Dev2: **9**
		Auto External Capture	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **10** • Head of Dev 2: **9**
		Auto Indexing - Sorting	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **9** • Head of Dev 2: **7**
		Auto Cleansing	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **10** • Head of Dev 2: **10**
	Data Management	Data Enrichment	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **6** • Head of Dev 2: **6**
		Data Governance	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **6** • Head of Dev 2: **6**
		Data Deduplication	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **9** • Head of Dev 2: **8**
		Data Cleansing	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **10** • Head of Dev 2: **10**
		Data Normalisation	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **8** • Head of Dev 2: **7**
	Data Access	Data Access Tools	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **9** • Head of Dev 2: **7**
		BI for MDM	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **9** • Head of Dev 2: **8**
		Internal Collaboration Tools	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **9** • Head of Dev 2: **9**
		External collaboration Tools	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **8** • Head of Dev 2: **9**
		Mobile Access	• Head of Dev 1: **1** • Head of Dev 2: **1**	• Head of Dev 1: **5** • Head of Dev 2: **5**

Business Transformation Program Director	Automation	Auto Internal Capture	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **6** • BTPD 2: **7** • BTPD 3: **7**
		Auto External Capture	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **6** • BTPD 2: **6** • BTPD 3: **5**
		Auto Indexing - Sorting	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **5**
		Auto Cleansing	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **5**
	Data Management	Data Enrichment	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **5**
		Data Governance	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **8** • BTPD 2: **9** • BTPD 3: **7**
		Data Deduplication	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **9** • BTPD 2: **9** • BTPD 3: **7**
		Data Cleansing	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **9** • BTPD 2: **9** • BTPD 3: **6**
		Data Normalisation	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **4**
	Data Access	Data Access Tools	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **5**
		BI for MDM	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **4**
		Internal Collaboration Tools	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **6**
		External collaboration Tools	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **5** • BTPD 3: **4**
		Mobile Access	• BTPD 1: **1** • BTPD 2: **1** • BTPD 3: **1**	• BTPD 1: **4** • BTPD 2: **4** • BTPD 3: **4**
IT Project Manager	Automation	Auto Internal Capture	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **9** • IT PM 2: **9**
		Auto External Capture	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **10** • IT PM 2: **10**
		Auto Indexing - Sorting	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **7** • IT PM 2: **9**
		Auto Cleansing	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **10** • IT PM 2: **10**

	Data Management	Data Enrichment	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **9** • IT PM 2: **7**
		Data Governance	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **10** • IT PM 2: **10**
		Data Deduplication	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **9** • IT PM 2: **9**
		Data Cleansing	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **10** • IT PM 2: **10**
		Data Normalisation	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **7** • IT PM 2: **7**
	Data Access	Data Access Tools	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **7** • IT PM 2: **6**
		BI for MDM	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **5** • IT PM 2: **5**
		Internal Collaboration Tools	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **6** • IT PM 2: **8**
		External collaboration Tools	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **6** • IT PM 2: **7**
		Mobile Access	• IT PM 1: **1** • IT PM 2: **1**	• IT PM 1: **4** • IT PM 2: **4**
Head of Business Analysis	Automation	Auto Internal Capture	• Head of BA: **1**	• Head of BA: **9**
		Auto External Capture	• Head of BA: **1**	• Head of BA: **7**
		Auto Indexing - Sorting	• Head of BA: **1**	• Head of BA: **5**
		Auto Cleansing	• Head of BA: **1**	• Head of BA: **9**
	Data Management	Data Enrichment	• Head of BA: **1**	• Head of BA: **5**
		Data Governance	• Head of BA: **1**	• Head of BA: **9**
		Data Deduplication	• Head of BA: **1**	• Head of BA: **9**
		Data Cleansing	• Head of BA: **1**	• Head of BA: **9**
		Data Normalisation	• Head of BA: **1**	• Head of BA: **5**
	Data Access	Data Access Tools	• Head of BA: **1**	• Head of BA: **4**
		BI for MDM	• Head of BA: **1**	• Head of BA: **4**
		Internal Collaboration Tools	• Head of BA: **1**	• Head of BA: **5**
		External collaboration Tools	• Head of BA: **1**	• Head of BA: **4**
		Mobile Access	• Head of BA: **1**	• Head of BA: **3**
Business Analysts	Automation	Auto Internal Capture	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **9** • BA 2: **9** • BA 3: **9** • BA 4: **10**
		Auto External Capture	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **7** • BA 2: **8** • BA 3: **8** • BA 4: **7**
		Auto Indexing - Sorting	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **4** • BA 2: **4** • BA 3: **5** • BA 4: **4**
		Auto Cleansing	• BA 1: **1** • BA 2: **1**	• BA 1: **10** • BA 2: **10**

			• BA 3: **1** • BA 4: **1**	• BA 3: **10** • BA 4: **10**
	Data Management	Data Enrichment	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **4** • BA 2: **5** • BA 3: **5** • BA 4: **5**
		Data Governance	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **10** • BA 2: **10** • BA 3: **10** • BA 4: **10**
		Data Deduplication	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **10** • BA 2: **10** • BA 3: **10** • BA 4: **10**
		Data Cleansing	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **10** • BA 2: **10** • BA 3: **10** • BA 4: **10**
		Data Normalisation	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **5** • BA 2: **4** • BA 3: **5** • BA 4: **5**
	Data Access	Data Access Tools	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **7** • BA 2: **5** • BA 3: **5** • BA 4: **6**
		BI for MDM	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **4** • BA 2: **5** • BA 3: **4** • BA 4: **5**
		Internal Collaboration Tools	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **4** • BA 2: **4** • BA 3: **4** • BA 4: **5**
		External collaboration Tools	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **4** • BA 2: **4** • BA 3: **4** • BA 4: **4**
		Mobile Access	• BA 1: **1** • BA 2: **1** • BA 3: **1** • BA 4: **1**	• BA 1: **4** • BA 2: **3** • BA 3: **3** • BA 4: **3**

Interview Question 28.

Explanation and reasoning for question 28.

Secure Senior Management Support	This question is designed to lead to a summary of the view that the interviewee has with regards to the support that the project gets from the senior management.
Implement a Formal MDM initiative	This question is designed to lead to a summary of the view that the interviewee has with regards to the way that the Master Data Management Project and Data Governance Policy has been planned from the early stages to the delivery and how satisfied the interviewees are with this plan.
Implement Standardized training	This question is designed to lead to a summary of the view that the interviewee has with regards to the training that has been designed for each part of end-users are taking part on the Master Data Management and Data Governance process. The interviewer wants to understand the level of mature thought and planning that has been put into the training process and how important that is based on the interviewees' opinion.
Invest in automation	With this question, the interviewer wants to discuss in a summarised way if the interviewees are comfortable with the level of automation that Master Data Management process is designed and how confident as well they are with the limited control over managing the golden record if they have any limitations.
Enable Remote Access	With this question, the interviewer wants to discuss in a summarised way if the interviewees are keen to have access to the Master Data Platform through remote access. A subsequent question would be to understand the reason that this is flagged as a need by the interviewee.

Appendix 2.

Company One Audit, Company Background

In October 2011, Former Group name of Company One decided to move forward with a complete transformation for both business and IT, by implementing a state of the art new Enterprise Resource Planning (ERP) system replacing the two legacy systems, which at that time were both summing up to a 50 years of age. The decision to implement the new ERP system was made just after a new state of the art implementation of a new Warehouse Management System (WMS) at a new Distribution Centre in West Yorkshire.

Since the successful implementation of the new WMS was based on Microsoft Technologies, along with the IT strategy at that time which was based in Microsoft Technologies, made the selection of the new ERP, an option that it had to be inline –technologically as well as strategically- with what the strategy was. Thus, the selection was Microsoft Dynamics AX which is the market leader.

Following the market and technological trends, and in addition to the nature of the business, data is the most valuable asset in an organisation and especially for a business-like Company One, the integrity of this asset is the most important factor for moving forward as a business and of course for generating revenue. An example of this need is the Vendor Rebates process.

Understanding the importance of this, the business decided to invest in a Master Data Management solution.

1. There are a large number and variety of systems in the Company One's technological ecosystem interspersed across every sector and department of the group. Products, Customers, Vendors and employees are the main datasets that each function uses to perform their tasks and services.
2. All this transformation that was mentioned above could not happen without populating the new ERP with a proper set of master records. For this reason, a clean, de-duped, "golden" version of the Customers, Products, Vendors and employee's dataset should be produced from all of the incumbent systems, which were in use within the ecosystem, and stored in a Master Data Management Solution that would ensure this uniqueness and single version of the truth, for every single record of these datasets.
3. Being a complex and growing organisation, Group's IT function would continually evolve, creating the need for further integrations with the introduction of new functions, processes and systems. This evolution brings along new Datasets that will need to be considered as part of the MDM strategy. Having already a predefined system that measures, manipulates and merges this data, is fundamental.

Company One Audit Method

The audit process started on Monday 2nd of March 2015 at 09:30 am at Company One Headquarters.

This is an external audit by the researcher and is used to review specific information relating to the MDM operation of Company One.

Preliminary Risk Assessment

The first stage of this audit was to undertake a preliminary risk assessment which included Interviews with the company's management. For this particular stage, the Interviews were carried by the researcher and the CIO of Company One.

In these interviews, Company One CIO established the Master Data Management process within the Group, as one of the highest Business Risk areas within the ERP Project (AX implementation) as well as within the business.

Planning Stage

The second stage of this audit was the planning stage, which included multiple interviews with key employees that run currently the MDM process. The conversations were based on the aims and objectives that came out from the first interview within the preliminary risk assessment stage.

The interviews were followed with Director of MDM Company One; Project Manager for MDM (an external contractor) and a Business Analyst for MDM (also an external contractor).

The forenamed were selected for an interview, based on the nature of their role Following these interviews, the researcher identified three more individuals, who were also interviewed: The Head of IT Development Company One, a Software Specialists and Database Developer and the Head of Enterprise Architecture Company One.

Testing Phase

The third stage of this audit was the Testing phase. During this stage, the researcher reviewed the business processes related to MDM to determine any potential violations of the operational standards or best practices, a set of documentation was reviewed and has been combined with the notes that were taken from the interviews.

Exit Meeting

The fourth and final stage of this audit represents the wrap-up phase of the audit methodology. In this meeting, the auditor and the company management will review the audit results and discuss any major violations or failures discovered during the testing phase as well as define a set of objectives and actions based on the findings.

Area	Document Type	File Name	File Type
Customer	Business Requirement Document	Company One brb Customer 2014-10-11 v1.0	Word Document
Customer	Business Requirement Document	APPENDIX A_User stories Customer 2014-11-18	Excel Document
Customer	Business Requirement Document	APPENDIX B New Account Creation for New and Existing Customer 2014-11-26	PDF Document

Customer	Business Requirement Document	APPENDIX C_NFR - Customer Data Validation 2014-11-06	Excel Document
Customer	Business Requirement Document	APPENDIX E_Stakeholder's List	Word Document
Customer	MDM Solution	Customer MDM Solution	PowerPoint Document
Customer	Process Maps AS-IS	AS-IS Business Acquisition 2014-06-25	PDF Document
Customer	Process Maps AS-IS	As-Is New Account for Existing Business 2014-10-29	PDF Document
Customer	Process Maps AS-IS	As-Is New Customer Creation (Corporate) 2014-08-05	PDF Document
Customer	Process Maps To Be	TO BE New Account Creation for New Customer 2015-02-18	PDF Document
Product	Product Information System	Business Requirements Document 2015-01-26_v1 1	Word Document
Product	Product Information System	Service Level Agreement xxx v1.0	Word Document
Product	Product Information System	TO BE End to End Product Creation (xxx) 15-03-02	Word Document
Product	Business Requirement Document	APPENDIX A_User stories Product Data 2014-07-24	Excel Document
Product	Business Requirement Document	APPENDIX B_TO BE Product Approval Process 2014-09-04	PDF Document
Product	Business Requirement Document	APPENDIX B_TO BE Product Approval Process 2014-10-17	PDF Document
Product	Business Requirement Document	APPENDIX B_TO BE Product Discontinued 2014-10-17	PDF Document
Product	Business Requirement Document	APPENDIX C_NFR 2014-02-21	Excel Document
Product	Business Requirement Document	APPENDIX D DRAFT ERP Product Data Dictionary v1.2	Excel Document

Product	Business Requirement Document	APPENDIX E_Stakeholder's List	Word Document
Product	MDM Solution	Responsibility Assignment Matrix	Excel Document
Product	Business Process Maps AS-IS	AS-IS Product creation 14-07-22	PDF Document
Product	Business Process Maps AS-IS	AS-IS Product Creation Specials	Visio Document
Product	Business Process Maps AS-IS	AS-IS Product Discontinued 2014-08-18	PDF Document
Product	Business Process Maps AS-IS	AS-IS Specials Product Creation 14-06-18	PDF Document
Product	Business Process Maps AS-IS	AS-IS Specials Product Creation 2014-06-18	Visio Document
Product	Business Process Maps AS-IS	EPC Process Map notes 1.0	Word Document
Product	Business Process Maps AS-IS	AS-IS EPC Process 2015-02-16	Visio Document
Product	Business Process Maps AS-IS	EPC Process Map notes 1.0	Word Document
Product	Business Process Maps TO BE	TO BE End to End Product Creation 15-03-03	Visio Document
Vendor	Business Process Maps TO BE	TO BE Vendor Process 2014-11-03	PDF Document
MDM Solution	Solution	Company One MDM Solution Design Document DRAFT v0 7 020714 (2)	PowerPoint Document
MDM Solution	Solution	Company One MDM Solution Design Document DRAFT v0 9 070814	PowerPoint Document
MDM Solution	Solution	Company One MDM Workstream Design Review	PowerPoint Document
MDM Solution	Architecture	Integration	PowerPoint Document

MDM Solution	Architecture	TO BE State	PowerPoint Document
MDM Solution	Framework	Company One ERP MDM and Data Migration Framework 060514	PowerPoint Document
MDM Solution	MDM Roadmap	MDM Roadmap	PowerPoint Document
MDM Solution	MDM Roadmap	DATA Quality Meeting	PowerPoint Document
MDM Solution	Critical Data	EASE IMPACT FOR BUSINESS CRITICAL DATA (ALL) (2) (2) (4)	Excel Document
MDM Solution	Data Dictionary	DRAFT Horizon Product Data Dictionary v1.3	Excel Document

Appendix 3.

Company Two Audit, Systems Definition:

Accounting Software: This software was part of ERP2 and was taking the data directly from ERP2. That said, all the necessary information for UDM could be found in ERP2. However, for specific reports with specific user access and exclusive functionality from UDM, the accounting software could be a part of the sources and could be investigated further.

Temporary Stock Management System: A temporary stock management solution that probably will be replaced by WMS1. Further investigation was required on the process flow as it was not clear at this stage.

CRM: CRM was a standard Microsoft CRM product. The auditor didn't perform any investigation on the CRM as he knew the system and he knows where and how the data were held. Once the UDM project will initiate, further discussions with Marketing and Retail Operations would have to take place to understand detailed requirements.

However, customer profiling and segmentation would be a very important part of the Unified Data Model. Extended functionality would be built on top of UDM to support this process for the Data Take On. To do that though, there was a need for further analysis of the transactional history of each customer and based on the data, a profile for each client would be created. The main usage of the current CRM would be the Customer Data and the creation of a 360o view of Company's Two Customers via Unified Data Model.

E-Commerce: This was the platform that operated the online shop. All the transactions for the online customers were stored within this platform and at that moment there was an integration with ERP1. At that stage, there was a specific definition of how this will work on ERP2 as this is under development (E-Commerce V3). However, one of the important things that I'll need to investigate is the data profile of the E-Commerce transactions and data with the ERP2 equivalent with regards, to Products mainly, customers and of course transactions.

ERP2: ERP2 is very comprehensive and easy to use the system. The functionality is rich and it takes short time to familiarise with the system. The back end database is Progress and the structures of the schema are very simple to read and understand. There is very limited access to the back end due to the restrictions from the vendor and the export facility is available in three different ways:

- File export
    - o Suggested by the vendor
    - o Not preferred option by the Company Two IT department and the auditor o Very limited Real-time export
- APIs
    - o There are standard APIs available but for any bespoke requirements, there will be a cost involved by the ERP2 Vendor. The auditor asked the vendor if there is a way for Company Two IT department to create an API and then pass it to the Vendor for review and deployment, the Vendor said that this option is not available.
    - o Very controlled Real-time export
- Linked Server via OpenEdge ODBC driver

- The most appropriate approach since APIs are of the huge cost of Vendor allows controlled usage of the linked Server functionality as they limit the number of queries that the Company Two IT department can perform
- The auditor will have to investigate why this is happening and how this can be resolved.
- All the above can be used and based on the availability that the vendor will offer to the company, the level of real-time integration will be defined.

The data structures however are well defined and the extraction of the required data for the UDM will not cause any issue or problem.

There is a discrepancy on the way that ERP2 handles the transactions as it creates three different levels on anything that has to do with the Order Processing (OP).

ERP2 uses a different hierarchy from all the other systems that are currently used in Company Two. It creates:

- *Order Header*
- *Order Lines*
- *Item matrix*

This wouldn't be a problem at all if all the other systems that are mentioned on the list of systems that have been investigated were using the same hierarchy with the above.

This is something that UDM can handle but this is something that needs to be looked at with the other systems as this implicates discrepancies with the product codes that the other systems are using.

Product Information Management System (PIMS2): PIMS2 is an extra module of ERP2 and is a dedicated Product Information Management system. The products are created there and they are taking a code which later is processed to ERP2. There isn't any integration issue between ERP2 and PIMS2 as they are both developed to talk to each other.

The data structures are well defined and there will be no issue for exporting the data into UDM. However, there are the same options for exporting with ERP2 and it has the same limitations.

There is a discrepancy here with regards to the product structure with the other systems. PIMS2 and ERP2 have a completely different product structure with all the other systems as of now. ERP2, ERP1, Retail and Shop system, WMS1 have completely different product structures with each other (Retail and Shop system and ERP1 share the same product structure as well as the current version of E-Commerce)

The biggest concern at this point is that the teams are working in separate silos not following a specific Data Strategy on how to follow a specific product creation policy. This will probably cause problems when the systems will go live but most importantly when the business will want to relate previous data with the new world data as they will have to do the same work multiple times based on the system. This is a part of Data Governance Policy and Data Retention Policy which from what the auditor have gathered so far there isn't any.

Shop and Retail System: This is the till Register system that is used on the stores. This system is a SaaS system with SQL Server database backend. The structures are simple but the database is in French. There is a process that exports files into an FTP location and any Company Two services, picks it up from there.

Static and Transactional data are well defined and the export-import routine on UDM will not be an issue. The system is connected with a scanner at each till programmed on AEN 13.

At the moment, the shop and retail system's data is synchronised with ERP1 which is a different structure from ERP2. Once ERP1 go live there needs to be a Data Strategy that defines the data conversion to the new data formats

Warehouse Management System (WMS1): WMS1 is the new WMS that will be linked directly with ERP2. The data structures are very well defined and due to the auditor's previous experiences with the product, there is a high level of familiarity with the data structures, the data format and the export routines.

There is a growing concern though with the content of the data within WMS1 as it has been identified that Product definitions are completely different from any other system including ERP2 and PIMS2.

However, the auditor hadn't identified yet if the content of the data that exists within WMS1 is just testing data or the actual data that is going to be used when ERP2 goes live.

There is also an irregularity on the definition of the product within WMS1 which if it is not due to testing data, it will cause a major issue in the business process in general.

Another growing concern with regards to WMS implementation is the alignment between ERP2 Release date and WMS1 release date. From previous experience, WMS1 to operate will require ERP2 to be the main data provider. Based on that, if ERP2 is not ready for release, The Company Two will have to put WMS1 aside and continue to operate the warehouse via ERP1 or Connect WMS1 with ERP1 which is a redundant task as ERP1 operates the stock management on its own.

At that moment, and this is based only on what the auditor has seen so far and only to the auditor's knowledge, the fact that there isn't an Enterprise Data Strategy and Data Governance policy, forces the different teams of the BTP to work in silos without following specific rules or guidance on data creation and data management. This will be a major issue if it will not be addressed at the early stages of the BTP.

ERP1: ERP1 at that moment is the main operational tool of the organisation. Every process run form ERP1 and every process from all the other systems ends up in ERP1. At that moment there is a full harmonisation between ERP1 and the surrounding systems and the integration is happening in a normal way apart of the Business as Usual operational problems that occur in a daily basis and are being trouble-shot by Company Two IT.

The ERP1 database is an Oracle database and the Export routine can happen in multiple ways. There are already existing APIs and in house developed applications that are linked to ERP1 and the same approach will be taken into consideration for the Unified Data Model.

Appendix 4.
Interview Answers in Grid views

Question 3

Role	Function	Answer
CIO	IT	Both IT Directors answered that the Quality of the Data is extremely important for them and their function. One of them mentioned that it is extremely hard to ensure that the quality of the data is kept to the highest level.
CFO	Business	All three Finance Directors focused on the importance of Data Quality on reports. One of them said: "It is very important to know that the report that I have in front of me is accurate and that I can trust it". The common part of their answer was the fact that they need to know the financial state of the business at any time so they can make decisions. Each process of the business needs to be constantly evaluated so the cost will always remain low and the profit will be optimum.
MDM Directors	IT/Business	As expected the Master Data Management director embraced the importance of Data Quality. As the role indicates ensuring that the Quality of the Data is at the highest standards is the Business as Usual action.
Head of Development	IT	Both Heads of Development said that data quality is very important as every application that is delivered is based on data exchange and data integration. If the data is of poor quality, the process that the applications are executing will fail and "IT will take the blame".

Business Transformation Program Director	Business	The Business Transformation Program Directors gave the interviewer the biggest surprise. Two of them recognised that the quality of the data is important but their "job is to make sure that the project will be delivered". The other one though said, "We are not so concerned with taking Data over from ERP1 (This the Company Two Program Director from the second audit) if it is inconsistent to start with". Also, two of the Program directors had Data Transition and Data Quality on their programme plan and the third one had the Data Transition and Data Quality as last part (after 14 months on the project) without any plans on Data Strategy in general.
IT Project Manager	IT	Both Project managers understood the importance of Data Quality and both said that to have "a successful project delivery you must ensure that your data is correct". When the interviewer asked them how do they ensure that the Data Quality requirements are met during the conversation, one of them answered: "This is a business activity" and that "the business needs to decide".
Head of Business Analysis	IT/Business	The head of Business analysis said that he understands that the data needs to be correct but their role is to ensure that they write down the "AS-IS" and "TO BE" process in time as requested by the Project Manager and not to do data analysis. When the interviewer asked them if their "AS-IS" and "TO BE" process definition includes data definition and rules for Data Integration, their answer from both of them was No.
Business Analysts	IT/Business	Out of the four business Analysts, only one said that the Data Quality analysis and the data definition before business process description is necessary because this is how he will define the process by "visualising the data journey". The other three said that the data (not mentioning the word quality) is important but they spent most of their time documenting what the end-user is doing and they don't do any data analysis as this has not been asked by their manager.

Question 4

Role	Function	Answer
CIO	IT	Both IT Directors said that they understand what Master Data is.
CFO	Business	All three CFOs said Yes.
MDM Directors	IT/Business	At that point, the interviewer thought that he shouldn't ask the Master Data Management Director if he is aware of the Classification "Master Data"
Head of Development	IT	Both Heads of IT Development said yes.
Business Transformation Program Director	Business	All three Business Transformation Program Directors said yes. (One of them asked "You mean the static Data?")
IT Project Manager	IT	Both IT Project Managers said yes.
Head of Business Analysis	IT/Business	He answered with a yes.
Business Analysts	IT/Business	Three answered with yes and one with no.

Question 5

Role	Function	Answer
CIO	IT	Both IT directors defined Master Data Management as "The tool that manages and maintains the Golden Record".
CFO	Business	The Financial Director's definition was: "A Clean version of Customers". When the interviewer asked them what do they mean by "Clean" the answers were mainly a de-duplicated version of the customer which enables accurate reporting.

MDM Directors	IT/Business	The MDM director defined as Master Data Management as “Managing a single version of the Truth”. As his role responsible for the data that goes to Customers/Dealers and on the website, his focus is more on the content of the products. That is why the MDM director’s definition was focusing on the product content.
Head of Development	IT	The Heads of development defined as Master Data everything static and that needs to be managed through a slowly changing dimension. By slowly changing dimensions the Heads defined the sets of Static Data that an update or deletion needs to be captured for accuracy with regards to reports.
Business Transformation Program Director	Business	All the business Transformation Program Directors defined Master Data as a “Clean set of Customers, Suppliers and Products”. The Program directors, focusing on Business Transformation were referring more to a Data Take On (Data Transition from the old systems and processes to the new) strategy.
IT Project Manager	IT	All the IT project managers defined as Master Data all the static data that are going to be used on their projects.
Head of Business Analysis	IT/Business	The Head of Business Analysis defined as Master Data the set of Customers, Products and Suppliers that are going under a profiling exercise to be cleaned.
Business Analysts	IT/Business	The business Analysis following from the Head of Business Analysis, defined as “Master Data the set of Customers, Products and Suppliers that are used for the process definition. This Data needs to be cleaned and de-duped.”

Question 6

Role	Function	Answer
CIO	IT	Both CIOs declared that Master Data Management is very important for any of their project because they cannot deliver projects where the static data are not maintained. Also, the only way to maintain the data is through a Data Governance Policy that has to be understood by the business. If the business will not follow a Data Governance policy that will force their employees to follow, the Master Data Management will never deliver the expected results and every project delivery that is undertaken by IT will always have problems leading to more cost on maintenance and manual non-standardised fixes of the problems.
CFO	Business	The CFOs had a different view on the importance of Master Data Management. They both focused on the accuracy of the reports and the cost of maintaining this accuracy. The cluster as the most important part of the Data Management (General not only Master) to keep their insights accurate and their cost low.

MDM Directors	IT/Business	According to the MDM Director, Master Data Management is very important. This process is "Business as Usual" for him and his team. "The content of the important attributes to define a product is important to be accurate as well as the single version of the truth for the products. My business process needs to ensure that that list of products that I promote to the business is clean and with no duplicate records. Once I sent the data to the business the list is used for many other operations. Except for the E-Commerce website that needs to be -spot on- I need to make sure that the products are clearly defined in the warehouses, so the business will know exactly what levels of stock they have at each time."
Head of Development	IT	Both Heads of Development understood the importance of Master Data Management and also embraced the importance of the process for their projects. At the moment they are facing problems with development that is due to bad quality of data. One of them said: "Poor Data in; Poor Data Out". Most of the times when they are trying to deliver a project the inconsistency between the data is so bad which takes most of the time to understand what they process in comparison with the time that they need to develop the application and deliver the project.
Business Transformation Program Director	Business	All the Business Transformation Program Directors understand the importance of Master Data on their Deliverables but it seems like that they don't plan for Master Data Management as a processor Maintaining Data Quality. They declare that it is more important to deliver the Program and define a Data Strategy after the delivery because it is difficult and time-consuming to go through the business users and ask them to take time out of their daily tasks to define the "AS-IS" and "TO BE" process as it stands. If they were to ask them about Governance and implementation of a Data Governance at the same time, the Program would have faced difficulties to be

		delivered on time. One out of the three interviewees though said “If the Data are of bad quality, we are not taking it into the “new world”. We are focusing on the new process and the old data can come later!”
IT Project Manager	IT	The project managers declared that Master Data Management is important, but if it doesn’t exist as a deliverable objective, they cannot do anything about it. The business needs to decide at the beginning to plan and budget a Master Data Management tool or process and if this is approved then they would initiate a plan that includes a Master Data Management Process.
Head of Business Analysis	IT/Business	The head of business analysis identifies as Master Data Management is an important part of the process, however, if this does not exist as a priority by the Project managers they cannot spend time on business analysis and analysis, in general, that is not part of the general roadmap and plan.
Business Analysts	IT/Business	The business Analysts said that the Data needs to be clean and tidy in order the processes that they design to be effective but they are more focused on defining the process as it is now and how it is going to be than planning on how to ensure that the quality of the data is the required. They assume that the quality of the Data and the Master Data, in general, is the desired unless the user during the process definition recording will mention a process that has to do with the Master Data Management and usage.

Question 7

Role	Function	Answer
CIO	IT	The IT Directors were not aware that a Master Data Management can be characterised in three different categories but they assumed that there would be something out there to define the Best in the Class method. They said that it would be ideal to reach the Best in class status but both they suggested that unfortunately at the moment they are have nothing in place to ensure Master Data Management Process and that they would be very happy to reach the Laggard Status as a first milestone.
CFO	Business	The CFOs were not aware either that a Master Data Management process can be classified and they never thought about "putting a label on the way that we are managing the Data" (Not specifically for Master Data but for Data in general).
MDM Directors	IT/Business	The MDM director said that he could classify the Master Data Management process with regards to Data Completeness and the level of "cleanliness" that the data is. However, he did not mention anything about the process in general. The only classification was with regards to the content of the Data and how accurate this data is.
Head of Development	IT	One of the Heads of Development said that the Best in Class classification should capture the data on creation or alteration and check on the Master Silo if an existing record exists or if the rules are applied to proceed with the transaction. However, the level of systems that the organisation has in place would never allow this kind of process to be implemented as an extra step before the transaction could be committed. The other Head of Development said that they are trying to capture the data on the creation and that they have a comparison process before the record been inserted on the destination however this process is in place only a small set of systems. Both of them seems to be aware of the classification of the process and they would put their current system to a lower level than Laggard.

Business Transformation Program Director	Business	None of the Business Transformation Program directors was aware of the classification that can be applied to the Master Data Management Process.
IT Project Manager	IT	One of the IT project managers as part of a project in a previous role that had to do with the creation of a Centre of Excellence for Data Management so he had come across the classification. However, his main focus was to manage the budget most profitably. The Centre of Excellent for Data Management was at the initial stage of establishment and he didn't have to get involved in the deep details. The other IT project manager was not aware.
Head of Business Analysis	IT/Business	The Head of Business Analysis was not aware of any classification with regards to Master Data Manager however he thought that there would be specific Key performance indicators that separate the processes from optimum to underperforming.
Business Analysts	IT/Business	None of the business Analysts was aware of any classification with regards to Master Data Management

Question 8

Role	Function	Answer
CIO	IT	Both CIOs declared that there is not a formal MDM system in place in their organisation and that there is not any budget assigned to implement one at that stage. Both of them measured the question with 1. Both of them said that they are tasking the IT to follow specific rules on the creation of the data on the systems that they use (SLA systems and incident management) and the systems that they deliver to the business but except very few occasions everything is manual. Within the Organisation the rules are easily overwritten and they are manual as well. The measured the question with an average 2. (The first one measured with 1 and the second with 3). For the third part of the question, the one CIO said that there is a collaboration with the MDM team -Content Creation using PIMS1- but there is not a collaboration with the other parts of the business. The other one said that there is very limited collaboration between the business and the IT and some of the times there are different points of view on what needs to happen. The average score for the third part was 2.5 as the one scored the action as 1 and the other as 4.
CFO	Business	All the CFOs said that there is not an MDM system in place and that they understand that this should be something vital that they can apply a form of Governance. All of them said that they should be the ones that they will promote and support this process and that they should ensure that a project like that should be on the near plans. Following that, the first part had an Average score of 1 as all the CFOs marked the question with 1. The same score for the second question however, all of them looked confident that the future systems will automatically force specific rules that the data capture will be handled and processed automatically. For the second question, they scored as 1 the current state. The future state though is got scored with an average of 6 having the one mark it with 8 the other one with 4 and the last one with 6. The latest score applies for the third question as all the CFOs looked pretty optimistic that the new world would enable better communication and collaboration between the business and IT. However, for the current state, the score was 4.66 as two of them scored it with 4 and one of them with 6.

MDM Directors	IT/Business	The MDM director said that there is a formal function of MDM but there is not a system in place. This is his next goal and this is his active project. He is going through a due diligence process with three vendors and then he has the costs he would take it to the board for approval. His final mark on the MDM system in place was 1 due to the current state. The second part which has to do with the automated capture of the data on the creation again he scored it with the mark of 1. The reason for that is even though they have a dedicated function within the organisation which deals with the Master Data, the business limits the liberties of this team only in managing the Product Data Domain. The creation of Customers or Suppliers is completely out of control and due to the state of the business which currently has 4 different ERP systems and 4 different CRM systems to operate, it is just unmanageable. Also, even though that the business has budgeted an MDM function, there is not a will to apply any governance on the creation of the records since they are "too busy" to deal with that. That leads to the next part of the question which again he scored it with 1 as the business has a group of people that deal with data due to the ERP consolidation project, but this group of people are not willing to listen to IT. They believe that this is the correct process because this is what they used to do with the old systems and that IT will have to deliver what they are asking for. They think that they know better what needs to be captured and how even though the IT suggests that this process is not optimal.
Head of Development	IT	All the heads of Development scored the first part of the question with 1. They said that there is not an MDM solution in place and they suggested they cannot put any KPI as even the processes on the new ERP implementation is not included in the overall solution design. The discussion went to the new world and how the business didn't consult the IT in principal to adjust the issue of Master Data. As a result, even though they are aware of the current situation that data from multiple domains are created in multiple different solutions, the new world will keep the same issue as Customers are still going to be created into CMR and ERP as different entities and Product can still be created into ERP as well as PIM system. That means that the score from all the Heads of IT Development was again 1 on the second part of the question. With regards to the third part of the question, the Heads of IT Development gave a score of 1 as

		they were suggesting that there is a disconnection between the IT the project team and the Development in general with regards to Master Data Management at that point.
Business Transformation Program Director	Business	All the business Transformation directors scored this part of the question with 1. The reason is that even if they are going through a complete change of process they don't have a separate system that acts as a Master Data Management solution. However, they were confident that the new system that they put in place will have all the required functionality to ensure "clean data". When the interviewer asked them how they said that it is a standard functionality of the system. The Interviewer went a bit further though and he mentioned the fact that on the new business Transformation ecosystem, multiple desperate systems allow the creation of multiple records under the same domain. Based on that how do they ensure that the creation or the alteration of the same information in different systems is aligned, one of them said: "that the new systems will take care of that process". With regards to the second part of this question, none of the three Program Directors could understand the reason behind the existence of this practice. When I explained them they were insisting that the "new world will not allow duplicates and will clean the data that will come from the other applications". Then the interviewer asked them what mark would they put on the current status and the future status of this process one of them scored 7 the other one scored 4 and the other one scored 6. At that point, the interviewer asked them how they can mark this high a process that does not and will not exist. They answered that "in the 'new world' the system will have a data cleansing process which will take care of the duplicates". On the third part of the question, all the interviewees answered that for each business function that has to deal with static data will be a member of staff that will be responsible for the "Static data". This maintenance will be happening through "the system" and if there is a specific request IT will be involved.

IT Project Manager	IT	On this question the IT Project managers marked the first part with 1 as on their projects there isn't any formal Master Data Management Solution in place. One of the two said that he asked about a Master Data Management solution for managing the static information but there weren't any immediate plans as at the moment the current project plan indicates that this needs to go under the "Data Take On" process with process-related individuals to validate the content of the static information through the system. The other IT project manager mentioned that on all the projects that he has worked the creation of a data dictionary for the statics was always taking the most of the time of the projects as most of the times there was a detachment between IT and business. IT could provide the attributes but the business couldn't validate the content and also that they were leaving the data mapping task at the end instead of setting it up at the beginning. That is why the other IT project manager suggested that a formal Master Data Management solution or "something that creates an index for the content between the business and IT should be in place" however both they would mark the formal Master Data Management solution with 3 as it is not in place, and when they realise the importance they start managing it manually. For the second part, they both stated that there is not any plan for automated data capture at the creation that is why they would mark the second part with 1. For the third part of the question, both project managers said that there is not a cross-functional team but the plan is that one dedicated person from the business function will be responsible for the management of the static data and when there is a problem will be going to IT.
Head of Business Analysis	IT/Business	The Head of Business Analysis suggested that there is not any formal Master Data Management Solution in place and "at the the moment the main focus is the definition of the processes". He suggested though that through the process there is a part that has to do with the data quality. When the interviewer asked the Head of Business Analysis to describe in more detail the part that Data Quality is involved, he mentioned that the Take on of Customer from the existing systems into a new the system needs to follow a cleansing process before insertion as well as Products and suppliers. When the interviewer asked him what would happen after the initial insertion, the Head of Business Analysis said that the new systems should be able to

		take care of any duplicates. Based on the fact that there was not any Master Data Management solution though in place, he marked the first part with 1. Moving forward to the second part of the question, the Head of Business Analysis described the current process of automatic capturing of the data as very important for the business but “at the moment is incomplete” there is not a validation of the data that is automatically captured and that in the new world, this process will still be with faults unless the business decides to follow a specific governance or policy. However, the systems that are going to replace the current ones will be in the position to manage the different versions of the updated data. With regards to data creation, the systems will be able to capture the data on the creation but will not be able to validate what has been inserted as the constraints within the system will be very limited and will be looking only for exact duplicates which in this case the system will not allow it. The mark that the Head of Business Analysis gave to that part of the question was 4. With regards to the third part of the question, the Head of The business analysis said that there will not be a cross-functional the team between the IT and the business to manage the data. The business will assign specific roles to specific users to act as data stewards in “potentially” a system that will be provided by IT, however, this system hasn’t been planned yet and these users will be managing the data via an excel. The mark, in this case, was again 4.
Business Analysts	IT/Business	All four of the business analysts marked the first part of the question with 1 as there is not any formal Master Data Management system in place and there is not any plan for any separate system as during that time, the main focus of the business is the delivery of the business transformation programme. Their answer to the second part of the question was varied between them as two of them marked the automated data capturing with 6 and one of them with 4 and the last of them with 3. The two analysts that gave the mark of 6 were supporting that the new process captures all the product, customer and supplier data on the creation through the 185 | P a g e system and that if there is a duplicate record within this system, the creation of a new record will not be allowed. However, when the interviewer asked them what happens on the other systems that the creation of the same

		item is allowed, they said that the process does not have a constraint on the other systems but only on the main ERP. The same conversation happened with the other two analysts and the same answer was given. The capturing and the check of the record is happening only on the system that the user tries to insert a record and it this control/check doesn't happen across the systems at the same time. However, all marked with 5 the third part of the question with regards to the cross-functional team between IT and Business. They said that there is not a cross-functional team to manage data but there is good communication between IT and the team that manages the data from a content/business perspective and whenever there is a problem IT will help them to resolve. Another synergy between these two teams that came up from one of the Analysts indicated that once the team that is responsible for the content which is maintained in an excel spreadsheet, they will pass this excel to IT to import it into a different system.

Question 9

Role	Function	Answer
CIO	IT	In this question, the interviewer found both CIOs aligned with regards to their data strategy. Their focus was on three things. Data Quality, ease of access and data archiving. Their data strategy had been structured against the fact that the information must be accurate, how easy this information can be accessed to get meaningful insights and how the archived data can be compared and analysed to give an understanding of the impact that inaccurate data can affect the business operations. Their data strategy should follow their business objectives to allow the organisation to get accurate insights so the business can make informed decisions.
CFO	Business	With regards to this question, all three CFOs had one common goal. This goal was the growth of the

		organisation. The goal is the same across all the CFOs and their need can be described in one word as “Accuracy”. All three of them want their strategy to focus on the accuracy so when they produce their teams produce the necessary reports the numbers to be correct. The problems that they have most of the time is “One report says £x and the same report that comes from a different department says £y”. The main problem that they want to resolve with the data strategy, is the elimination of different silos within the organisation that manipulate data. There should be an Enterprise-wide data strategy which enhances the links between divisions and create consolidated business insights.
MDM Directors	IT/Business	The MDM had a different approach with regards to Data Strategy in comparison to the CIOs and the CFOs. His main focus was around “People, Process, Technology”. He described his Strategy as a holistic view of the data across the organisation, but the most important part of the strategy is the Data Governance. He mentioned that to create a Data Strategy, the organisation needs to follow specific rules and most importantly to comply with these rules. As long as the rules are followed by the end-users within each department, the data strategy is a valid process. By the time that the policy is over-ruled, the strategy does not affect and it is not valid.
Head of Development	IT	Both Heads of Development suggested that their main point with regards to Data Strategy is all the applications and integrations that are creating to have embedded rules that are aligned with the process. Their main concern though is that most of the times the business assumes that these rules are known to the IT and when a problem arises, the IT has to re-focus on fixing the problem without that ensuring though that the process has been explained correctly. The data strategy based on their opinion should describe in detail all the rules that a process requires to avoid any misconception in any application or integration development.

Business Transformation Program Director	Business	The business Transformation directors had a different view with regards to the Data Strategy. They all thought that the Business should provide the Transformation programme with a Data Strategy and that their responsibility should be to apply it within their Program. With regards to how important is the existence of a Data Strategy or specific rules or policies with regards to data on their objectives, they all said that they understand the importance of the rules with regards to data but these rules need to come from the business. When the interviewer asked them about the current level or Data Governance within their programme, one of them said this something that will come from the software vendor and will be reviewed by the business in a later stage. Then the interviewer asked them about the way that they manage their releases on software if they don't have specific rules on how the data will be treated on each process, the answer was "we are more focused on the way that the new system should work in the way that the users are supposed to use it". When the interviewer asked them about the case that the users make a mistake or don't follow the correct flow that the system is designed to follow, the answer was "the system will be delivered in a way that will not allow users to use it differently than the way that it is supposed to work.
IT Project Manager	IT	Both IT project managers said that they didn't have any documentation with a specific Data Strategy. The only rules that they had to follow were the process within each IT deliverable that they had to follow. Their focus on each project is basically understand the input, the transformation if there is any and the output. This process definition though is something that they take from the Business Analysts and as long as the business analysts have defined the "AS-IS" process and the "TO BE" process it will be included on the project deliverable.

Head of Business Analysis	IT/Business	The Head of Business Analysis was actually in line with what the IT Project Managers were saying. The Head described that there isn't a specific Data Strategy and that the only aspect that they take into consideration with regards to the Data Management within each deliverable is the "AS-IS" and the "TO BE" process and how the data that is involved within the process are defined on the input and the output.
Business Analysts	IT/Business	The Business Analysts followed the same approach with the Head of Business Analysis and the IT project managers which is they translate the "AS-IS" and the "TO BE" process. What they noted though was the fact that
		when a business user explains them the "AS-IS" process and how this process is going to be altered to the "TO BE" process, they don't follow specific rules and that the "TO BE" process doesn't follow a guideline based on a Data Strategy. It is also mentioned that in most cases, the business users disagree with the fact that the new process is there to optimise the current process, as a result, the "TO BE" process to be a more complicated version of the "AS-IS" without taking into consideration any new data validations or rules that the users need to follow.

Question 10

Role	Function	Answer
CIO	IT	For this question, the CIOs' answer was interesting since the role of a CIO is a senior role. However, both of the CIOs said that they support the direction of a Master Data Management solution, however they find it "difficult to technically convince the senior management of the other business functions." For example, one of the two CIOs said: "All the board understands that there is a risk with the data within the organisation but the budget approval process for a project like that is very difficult." When it comes to budget CAPEX (capital expenditure) for a project that involves various functions of the business, there is always a debate on which business function should take the cost. This is when all the projects start reducing the feasibility. "In this organisation, the CAPEX approval committee meet once a month to approve projects that the estimated cost is above £x thousands. Usually, the operational challenges of the organisation go as P1 (priority one) and at the end of the appointment, a PowerPoint presentation needs to take place for each project. If this presentation is too technical, they will request a revision and a presentation on the next month's committee meeting. The next month, if the revised version of the presentation is simple and lack of details they will request again more details for the next month's meeting. At that time there have been 3 months trying to convince the board that they need the project based on a PowerPoint presentation as well as a usually 3 months' work before the first presentation for analysing the environment the requirements as well as the costs. At that point, we have spent 6 months trying to convince the board that they need to step up and support a project like that bearing in mind that all the challenges that are discussed on each meeting have to do with data quality."

CFO	Business	All the CFOs have agreed that they offer all the support that it is required on projects that will help the organisation to make better decisions and improve the business insights as well as the operational activities. One CFO mentioned that "I always support improvement on the process if this improvement is going to reduce the time, effort and cost of the current process without any quality reductions on the outcome of the process."
MDM Directors	IT/Business	The MDM director supported in a very direct way that the Senior Management does not support in practice the Master Data Management improvement. He suggested that even though he has a very large team that they are supposed to do Master Data Management, in reality, what they do is content management via excel. They don't use any sophisticated tools to help them understand and identify any incorrect records and they are cannot force a Data Governance policy across the organisation. The MDM Director suggested that every time he goes to the commercial function of the organisation to suggest improvements on the way that new customers are created, the commercial function replies with the fact that they understand that the process that they have in place is not correct but they don't have time to change it. "They are happy to talk about it but not change it".
Head of Development	IT	Both Heads of Development said that the Senior Management does not support the Master Data Management in practice but they support it in theory. The senior management realise that most of the challenges in an operational level come from a bad quality data however when the discussion comes around the implementation of a solution that would reduce or resolve these issues, there is a big difference between what they discussed and what they willing to

		do. Even though on any piece of application that they deliver as IT Development department, they are trying to maintain a high-quality data, they cannot as it stands at the moment to apply constraints on existing systems that limit the users' ability to enter inaccurate information on these systems. They cannot prevent the "Bad Data In – Bad Data Out" habit that existing systems allow at the moment.
Business Transformation Program Director	Business	On the other hand, the Business Transformation Program Directors suggested that the senior management supports the Master Data Management improvement that is why they invest in new systems and they redesign/transform the way that the business operates. However, when the interviewer asked them about previous conversations when they mentioned that Master Data Management is not part of their current project as the data quality and the management of the static information is handled by the new Systems, they mentioned that a Master Data Management system is not in the scope at the moment but the maintenance of Static information that is going to be used for any operational process is something important and that the Senior Management supports this process.
IT Project Manager	IT	Both IT project managers described the level of support from Senior Management as inconsistent. Within the multiple functions of an organisation you have some directors that support the Master Data Management as a process and as a "Must have" but there are others that don't understand the value. Usually, they end up with mixed messages in which they support the process but they cannot follow it.

Head of Business Analysis	IT/Business	The Head of Business Analysis suggested that the support from the business can be measured on the fact that a project to improve a process starts and requires analysis and definition of the business process. The implementation of this process always depended on correct and accurate data. The fact though that there is not a specific Data Governance in place and that there are no constraints in place to guide the user into entering the correct information or allowing the user to enter incorrect values without a specific valid meaningful information is since the senior management is not prepared to follow their own rules.
Business Analysts	IT/Business	The business analysts suggested that there is support from the senior management but there are always time constraints which don't allow the suggested plan to completed and followed as it has been scheduled at the beginning.

Question 11

Role	Function	Answer
CIO	IT	The solution that both CIOs suggested as the current one for managing and maintaining Data Consistency is a centralised Data Warehouse that integrates with all the systems. Within this data warehouse, the dimensions that define the schema of the Data Warehouse usually holds the Static information and within the dimension tables, there is a process that can Identify duplicate records. However, this means that transactions have already been made with inaccurate data and then through manual intervention by the developers and hand-coded rules by the developers these inaccuracies are usually are eliminated for any reporting. Once these inaccuracies are identified, the IT will inform the business function to manually adjust the records that contain inaccurate or duplicate data.
CFO	Business	All 3 CFOs said that the main problem that they have is how they recognise the payments that they have received and how they recognise which supplier they have to pay. This is something that at that stage happens manually by the dedicated departments' accounts payable and payment collectors. However, the reporting is happening by Business intelligence tools that are attached to the Data Warehouse database that is being maintained by IT.
MDM Directors	IT/Business	The MDM director said that the only tool that his team uses at the moment is an excel spreadsheet per data domain. Once this spreadsheet has a new version is sent to IT where IT uploads it to the relevant systems and the Data Warehouse. There is nothing sophisticated as it stands at the moment with the current process.
Head of Development	IT	With regards to the tools that are used to maintain a data consistency except for the integrations that are happening between the systems, they receive an excel spreadsheet with static information that will need to be imported after has been evaluated by the business. There is a process there that executes some tasks which process all the data to each system and the main data warehouse. If there is a duplicate, then the development team will have to understand the reason why this duplicate exists. Most of the times different

		functions of the business direct the IT team that there are duplicates for a reason. This reason according to IT is not valid but they cannot go back to the Business telling them that there shouldn't be a duplicate. Except for the automated integrations between the systems and the Data Warehouse, everything else is the manual investigation. The duplicates are identified only if there is a perfect match on a combination of attributes.
Business Transformation Program Director	Business	The Business Transformation Project Directors suggested that the new environments are going to support data consistency maintenance as a process within the systems. However, all the existing data that needs to be imported into the new systems will have to come from IT in an excel format. This excel will have to be manually validated by the process owners before inserted into the new systems.
IT Project Manager	IT	At the moment there is nothing sophisticated in maintaining data consistency across the systems. Both IT project managers agreed that everything is manual and is maintained in an excel spreadsheet. Once this excel spreadsheet has a final version is distributed to IT which will import it into the relevant systems.
Head of Business Analysis	IT/Business	The Head of Business Analysis said that at the moment is very difficult to maintain, manage and ensure Data Consistency across the systems. Everything is happening manually on an excel spreadsheet that is going to be used as the master file of static data to be inserted on the relevant systems. However, once this set of files will be inserted into the new system, the designed future process will maintain the accuracy of this data based on the process that exists by default for the static information.
Business Analysts	IT/Business	All the business analysts suggested that Excel spreadsheets are the tool that is used for all the static data. Nothing 193 | P a g e sophisticated and nothing complicated. Each Data Domain owner is responsible for maintaining this master file and once it is validated and confirmed, it will be distributed and uploaded by IT.

Question 12

Role	Function	Answer
CIO	IT	Implementing a Master Data Management solution is a very complex and costly exercise in which will give many benefits once the process reaches a specific maturity level. Both CIOs suggested that the cost that involves around an MDM process are massive and that because it requires CAPEX approval for every new software or hardware that is required to be purchased, OPEX approval for all the people that are going to be involved and dedicate time on this process. There must be a sequence of activities for any business transformation project and this sequence of activities need to start with a centralised Master Data Management solution. "What business does not understand though is the fact that work that is happening in other future projects with regards to Data Take On would have been eliminated if a Master Data Management process was in place"
CFO	Business	The CFOs' answers were on the same page with the CIOs' but slightly different. All three suggested that it will be a very expensive process at the beginning with regards to CAPEX and OPEX approvals and justifications but they think that the most expensive part will be the change of attitude within the organisations. People were used to doing tasks in a specific way and this transition will have to be as "smooth" as possible. What scares them more is ensuring the business continuity and the adoption of the new process from the end-users.
MDM Directors	IT/Business	The approach that the MDM director took in this question was a bit different from the previous interviewees. He mainly focused on the development of a Master Data Management solution as an in-house development or an "Off the shelf" product. He said that implementing the solution would cost quite a lot of the budget to be spent on hardware, software and development as well as changes 194 | P a g e on the existing systems to apply the new constraints which are driven by the Data Governance policy which is part of the Master Data Management solution. However, he said that with a "proper MDM tool" in place, the time that his team spent on manually managing the content of the data would be reduced enabling the team to focus on more productive tasks and moving forward reducing the number of the team

		members since the checking would now be happening automatically instead of manually. He suggested that with an MDM tool in the place he could reduce the team force by at least 50%.

Head of Development	IT	The Heads of development suggested that the cost that is involved in a Master Data Management implementation is massive for their department and they think that the main cost goes to them. They suggested that their team will have to be divided in a team that will take the responsibility to deliver all the task that are related with the implementation and a team that will have to work on a Business as Usual process as well as delivering projects outside the BAU and the MDM. Then once the MDM will reach a specific level, all these three teams will have to work together to manage the changes on the current systems which will integrate with the MDM. Then separate testing will have to be involved between these three teams as well as the business to ensure that the integration between the MDM and the existing systems is working as it should work.
Business Transformation Program Director	Business	The Program Directors of Business Transformation said that the cost would be huge based on a comparison that they did with the cost of their current project. A Master Data Management Implementation would involve all this analysis that has already be done to happen again to include the Master Data Management process and Data Governance. However, all three were pretty confident that the investment wouldn't add any value once the new systems are in place due to the capability of these systems.
IT Project Manager	IT	Both IT project managers have suggested that the cost for implementing a Master Data Management solution would be massive based on the size of the organisation and based on how many functions of the organisation are going to get involved. Another parameter that one of them added as well as the geography of the organisation as an

		the organisation that has different functions in a different geographical location would add more cost to the project's CAPEX and OPEX. They both understand the level of benefits that a process like that would add to the business based on the fact that tasks that they are dealing in a project like Data Cleansing would not be part of any new project.
Head of Business Analysis	IT/Business	From a Business analysis perspective, the Head of Business Analysis suggested that the cost would be quite big based on the fact that discussions about processes and training for these processes should take place before any implementation. They understand that the Return on Investment would be great but he suggested that based on experience, the business does not usually welcome a new "way of doing things" especially when a new version that produces the same output is introduced.
Business Analysts	IT/Business	Based on the current experience and the time that it takes for an organisation to make a decision, the business analysts suggested that the cost would be high. They suggested that it takes more time for the organisation to decide than move forward with the implementation of any project. Especially when we are talking about a project that involves both IT and Business and especially changes the behaviour of the business.

Question 13

Role	Function	Answer
CIO	IT	The answers to this question were different between the two CIOs. The first CIO suggested that every time he asks for a report from the IT department, he usually gets it within a couple of hours. However, when other functions of the business require reports, IT tries to have an SLA of 1 week before the delivery to the function that requested this report. Once the report is created then everything is refreshable. The second CIO suggested that he never thought it that way though. Usually, the problem lays in the lack of a data dictionary. It takes more time to create a data dictionary that has an explanation of the content of the data that the users usually request on a report than actually doing the report. The problem usually is a result of people deciding new rules on the process, they don't inform IT and when IT processes the data to generate the reports the data don't match.

CFO	Business	All the CFOs are not aware of this level of details however they all have weekly, monthly, quarterly and yearly reports. They know that each week the people that are preparing the report, they start 3 days before the requested day. 1 week for the monthly report, ongoing amendments on the quarterly report and about a month of reconciling the monthly reports with the yearly report because most of the time these reports don't give the same insights.
MDM Directors	IT/Business	The MDM director couldn't give a number with regards to this question. His team constantly receives data or directions from different functions of the team to manage the content and their job is to review the data that they receive every day.
Head of Development	IT	The Heads of Development suggested that the most difficult part is to understand where the business needs the data from, how the data link together and what the content means. One of them said: "there are so many different statuses for example for a stock movement that when we want to do a reconciliation of the inventory report, we have new statuses every time, as a result, to go back to the business to ask about this change and then wait for them to come back to us. By the time that they come back to us, we have moved to other Business as Usual requirements that have come as a priority. Then we go back to the report we take into consideration the new changes that we had asked for and when we amend these changes there is another status that came up."
Business Transformation Program Director	Business	The Program Directors of Business Transformation were not aware of the specific metric, however, they all suggested that with the new systems the reach for information should be very easy. When the interviewer asked them how much would you think that it would take for a user to find what they want they suggested that "The user should follow the process that is defined on the system. If the users use the system, the way that they are supposed to use it then they shouldn't have a problem."
IT Project Manager	IT	Both IT project managers described the search for

		information as “a nightmare”. They are not aware of how much time it takes to find the information that they need but both of them said that creating a data dictionary is the most difficult part of any of their projects since each time they have to start from the beginning. One of the said: “In one of the projects, we had as a task to create a data dictionary. The project lifecycle was 9 months and the attributes data dictionary took 3 to 6 months and the content data dictionary took 6 months. Even after the delivery of the application, the data dictionary was still not complete.”
Head of Business Analysis	IT/Business	The Head of business analysis wasn’t aware of these metrics either. His focus was more on the definition of the data content with regards to static information and he described it as a very difficult task. The users’ even if the content of the data is generated within their business function, most of the time they know only 40% to 60%. This makes it very difficult for his team of business analysts because most of the times they exclude from the process-specific parameters since nobody knows what they are used for.
Business Analysts	IT/Business	Following the answer of the Head of Business Analysis, the business analysts described the process of identifying the data a very difficult and stressful process since most of the times people are not aware of specific content that is very important for the process. Most of the time they have to deal with altering systems that they are really old and the definition of the “Statuses for example in an order” is not clear. “Users are aware of 5 statuses but when the IT comes back with 28 statuses then there is a problem”.

Question 14

Role	Function	Answer
CIO	IT	Both CIOs answered this question in a very simple way. They suggested that every amount of time that is saved on searching data would be a massive benefit for the organisation especially when this is a process that shouldn't take time in the first place. However, they suggested that each organisation is different and faces its own challenges. The information though no matter the nature of the organisation should be available instantly and the people who spent more time than what they need should learn to do their job in a better way. One of the two suggested: "There are some cases that people do the same job for 15 years. This is the only way that they know, these people's tasks are aware only to specific people within the organisation, as a result, IT not to be aware of the exact process. When these people doing something for 15 years and more in a dated technology it is their fault that they haven't shared or asked IT of what they are doing. These occasions would probably create false metrics".
CFO	Business	On this question, CFOs came up with a more strategic answer. They said that nine weeks is nearly a quarter if you calculate holidays and absence of the personnel that means that the best in class organisations concerning Data Management they spent almost a quarter of year less than the other organisation on searching which means that all the other organisations including theirs spent more. This metric triggered questions on CFOs like "How much does it cost in our organisation the search for information and how much does it take to produce the information?" Also, they said that if for example, their organisation spent twice the time on searching the information, that means that their insights are delayed so any decisions that they should have made in a specific time which could have generated more revenue to the business are potentially out of date. Except then from the fact that their employees have less time to do their actual job the insights are coming delayed and important decisions that they should have made in an earlier stage based on that metric they are left behind.

MDM Directors	IT/Business	The MDM director suggested that this number doesn't surprise him especially if he compares this number with the time that his time takes to go through the content of all the data that they have to manage it is very clear that the money that the business could have to save is enormous. First of all, due to deadlines that the business has to meet with customers that they receive data from the organisation, that would have eliminated the overtimes and weekend workings. Also, the fact that there are penalties for incorrect data that is sent to the customers this would have eliminated that as well. The other problem that is depended on specific timeframes is the rebates calculations, when the calculations are wrong that means that the supplier will not pay the rebates if they will not receive the information in time and accurate. This would have a massive impact on the business since the Rebates are a really big part of the annual EBITDA.

Head of Development	IT	The Heads of Development suggested by taking this time back, the team of the developers would be able to dedicate in other critical projects and deliver on time any other applications or integrations that are set back due to lack of data definitions. This time would speed up the process of systems delivery because the developers would have one single source that they trust to take the data from and based on that one big part of their development lifecycle would be reduced.
Business Transformation Program Director	Business	The business Transformation Program directors suggested that this would help their project delivery times as the biggest difficulty that they have at the moment is the data mapping and the data definitions which take most of the time. These 9 weeks could be used in different tasks reducing the cost of the Transformation project and also delivering more accurate and meaningful testing.
IT Project Manager	IT	The IT project managers following the same approach like all the others suggested that if they could reduce the time that it takes to do the Data mapping and the data definitions as part of each project, then the percentage of success rate would be higher from what it is so far. All the projects involve data within IT and usually, the definition is always different and where to find this data is always difficult.

Head of Business Analysis	IT/Business	The Head of Business Analysis answered this question with reservations. He said though that if the relevant people knew the meaning of the information that they are dealing every day and that they were not just doing mechanically a set of specific tasks, that would help any projects that the processes require analysis. Understanding the information that the users are dealing with, it would have always sped things up and the definitions of the process would be easier and more accurate.
Business Analysts	IT/Business	The business analysts suggested nearly the same thing with the Head of Business Analysis. They said that they spent so much time to understand what the data means from the users that they create it and uses it daily then go there and discuss the process with them.

Question 15

Role	Function	Answer
CIO	IT	Both CIOs said that each data domain should be maintained in a master source system. The customer usually should be mastered within a CRM, the products within a Product information management, the Suppliers within the ERP along with the Employees domain, the warehouse and the stock location should be maintained within a Warehouse Management system. The problem starts with the lack of discipline on the systems where customers can be created on the ERPs (when the business has more than one ERP systems), e-commerce platform, CRM, accounting software (potentially), EPOS systems etc. The product can be created everywhere and the supplier on the ERP, the CRM and PIM system. Usually, the actions that are in place are not taken under consideration as the end-users will do whatever it takes to do their job as fast as possible so this lack of governance and constraints allows these problems to exist.
CFO	Business	The CFOs suggested that the ERP should be the main platform that manages the creation of new data and that all the other system should take data from the ERP. Specific governance should be applied to the ERP processes to limit the cases that people create data everywhere.
MDM Directors	IT/Business	The MDM director suggested that each Data Domain should have a main system Master Source system that manages the domain. Once the data is managed in one agreed location, the governance and the maintenance is getting easier and the possibility for an error is reduced.
Head of Development	IT	The Heads of development suggested that the main problem that exists is the creation of data all over the data ecosystem. The end-users are not educated and they are not willing to follow specific directions on how to add and create new data. What the IT uses as the main source is the Data Warehouse in which potential duplicates can be identified within the dimension tables. Even if this method is not correct or it is not suggested, the people will always try to find ways to do reduce the list of tasks that they have in a daily basis by finding the parallel process that is not supposed to allow data creation.

Business Transformation Program Director	Business	The business Transformation Program directors suggested that the ERP should be the main system that is used as the source of data creation. All the processes are linked back to the ERP and from the ERP start all the processes that trigger transactions. Of course, other systems create records but the ERP should validate these records before processing them. Like for example the e-commerce platform or the EPOS systems. These two systems both create transactions with the majority of these transactions to be associated with a new customer. These new records of customers are integrated with ERP. At that stage, ERP should validate the creation of this new record and link it back to the e-commerce platform with a reference to the record that got created in the ERP.
IT Project Manager	IT	The most common problem according to the IT Project managers with regards to the Data Sources and how to identify which source should be the most important is the fact that all the sources that allow CRUD are all clustered by their users as important. It takes time to change a process that a business function is using for a long time and has been established as the preferred method of completing a specific process. When a new process requires discipline on the way that data is inserted or edited, it usually causes arguments between departments.
Head of Business Analysis	IT/Business	The Head of Business Analysis suggested that the ERP should be the main system that the control around the data is happening. All the governance should be applied there and since this is the most important system within the organisation and everything is depended on this, there should be more rules applied to the processes.
Business Analysts	IT/Business	The Business Analysts suggested that the ERP should be the main master source that holds the most accurate information and that all the information that is distributed to any other systems should be the one that comes from the ERP. Also, during the process definition, there have
		been many data dictionaries that have mapped all the necessary attributes for a process that involves more than one systems. This mapping should start from the ERP and should end back to the ERP.

Question 16

Role	Function	Answer
CIO	IT	Both CIOs suggested that all these problems that are mentioned on the question are problems that they have come across during the implementation of every data-related project. There is always the problem that it is described in question 15 where data are stored in too many different systems and every system lacks governance on creation edit and delete of the data. Of course, the time is critical since everyday costs depending on the number of people that are associated with each project, and that it also includes the time that is spent before any decisions for starting a project. And lastly, the manual data management on excel spreadsheets. Of course, there are other aspects as well, like agreements between the functions of the organisation that this project is required and why, the rise of CAPEX for each project and how this CAPEX is going to be converted later into an OPEX, how the budgeting will be affected for each year within the IT department for managing the deliverables of the project and so many other things that cause difficulties before the project start.
CFO	Business	The CFOs' main consideration was around time and cost which for them, time equates with the cost. The difficulties that the CFOs' usually have is the justification of the initial cost and the budget allocation. Also, the costs of hardware, software and project time involving employees and external contractors is a difficulty that usually on any initial project approval discussion is not represent the accurate numbers of costings, as a result later to try to find a way to cover the extra costs. On top of the above, due to the natural delays within a project, any extended time of a project has a domino effect on many other costs.
MDM Directors	IT/Business	The MDM director suggested that all of the above are difficulties for any data related project. For his vision of implementing his Data Strategy and his Master Data Management process though, he finds as the main difficulty the time that he and his team spent in a daily basis dealing with tasks that are not related to Master Data Management but on Business as Usual content management. Also, the fact that he has to repeatedly prove to the board of directors the

		value of the Master Data Management and the benefits that a project like this would give the business is very difficult. One of the most important reasons that the MDM director said that the implementation of a Master Data Management solution is that the "Master Data is too technical for non-technical people and too simple for technical people" to understand what is required to be done and how. The MDM director suggested that he spends more time trying to convince people on the senior management that they need to invest in a solution than actually planning the strategy for what his role is supposed to deliver.

Head of Development	IT	Both Heads of Development suggested that all the above are the main difficulties that they face during any data related project. However, again the most important difficulty that they face is the analysis of the process that each project needs to deliver. They suggested that it takes more time to understand what needs to be done and to agree that the final process is accepted between the related functions than actually delivering the project. The other difficulty that follows the one that just described is the time that is spent on mapping and defining data.
Business Transformation Program Director	Business	The Business Transformation Program Directors suggested that the most difficult tasks on any their project is time due to the nature of the business which is heavily based on time constraints and the initial data collection, definition and mapping manually. However, they all suggested that time is the most critical parameter.
IT Project Manager	IT	Both IT project managers suggested that a combination of all four difficulties that the interviewer suggested are the most common and from their perspective time is the most difficult constraint of the project to achieve. On a project that involves multiple data silos, multiple business functions and multiple different processes, it is very difficult to bring anyone "under the same page" and complete the objectives in the time that is required.

Head of Business Analysis	IT/Business	The Head of business analysis suggested that currently on the project that he works, he faces all the difficulties that are mentioned on the question. Trying to map multiple data between multiple different systems where the knowledge for some of the systems is limited it can be very difficult, time-consuming and with many risks. Then all these mappings between the systems are maintained in multiple large and difficult to review and navigate excel spreadsheets. And all of these actions are actions of the initial analysis. Concluding, he suggested that before the project even starts the development phase, the costs are already high.
Business Analysts	IT/Business	The business analysts similarly answered this question with the Head of business analysis. The suggested that all these 4 difficulties are visible on their current project and that the most difficult constraint is the time. Having to speak with multiple business users, is proven to be very difficult as they usually have to reschedule meetings to go through the process since the business users don't turn up.

Question 17

Question 17 was a supplementary question to question 16. The answers on question 16 covers this question as well.

Question 18

Role	Function	Answer
CIO	IT	Both CIOs said with no hesitation said that they would buy an off the shelf product instead of building an in house one. They both understand that there must be a lot of work that will need to be done before the implementation of the project as well as a lot of preparation work, but when the project and especially the environment reaches the level that is ready for Master Data Management, there will be more potential to deliver the actual value of Master Data Management than actually building a Master Data Management solution in house. One of them said, "There is no reason to re-invent the wheel". All the potential vendors that the CIO would go for, would pass through a due-diligence process. Introductory meetings will have to take

		place and based on the outcomes of these meetings would decide how they would move forward.

CFO	Business	The CFOs said that they would go with an of the shelf solution as well. The reason for that is that they can manage the deliveries better when the conversations and the deliverables are managed within a contract with specific breakpoints. The in-house development usually gets distracted by Business as usual as tasks that they have to deal with. Also one of the CFOs said that for an important project like Master Data Management, it is always good to have an external vendor to deal with the process as the level of communication between the in-house team and the external team starts from the beginning without any previous history which unfortunately most of the times on the in-house developed projects there is a history between departments and individuals that make the project journey more complicated.
MDM Directors	IT/Business	The MDM director, on the other hand, suggested that he would prefer an in-house Master Data Management solution since according to his sayings, he has done it successfully in the past. He understands the process and he understands what needs to be done. Also, it is his strategy that he wants to implement that is why he wouldn't like an external provider to offer a tool that could change his vision.

Head of Development	IT	The Heads of Development suggested that an of the shelf solution would be more beneficial for the business since the developers don't have the required experience to move to the next level of detailed and special algorithms. The organisations that the Heads of Development are part of, are not Software Houses therefor the teams have been the build-up to deliver solutions that will help the organisation to complete specific processes as well as to support existing ones. Producing sophisticated similarity algorithms and matching mechanism is not a skill that exists in-house at the moment and there shouldn't be any plan to develop something like that in the future. However, what they are specialised on is the Integrations and they would help a lot the Master Data Management process with all the integrations that would be required.
Business Transformation Program Director	Business	The Heads of business Transformation suggested that they would go for an off-the-shelf solution without a second thought.
IT Project Manager	IT	The Project Managers suggested that the off-the-shelf solution would be the most appropriate. The vendor of the Master Data Management solution has dedicated a Research and Development department only focused on an industrialised solution that works. The IT department with a massive list of business as usual tasks to do daily cannot turn into R&D overnight.
Head of Business Analysis	IT/Business	The Head of Business Analysis suggested that he would go with an off-the-shelf solution. He suggested that the vendor will have specific scripts to follow for the implementation as well as a specific set of tasks in a specific order. The Business Analysts of his team wouldn't have the required experience in defining the Master Data Management Process. That is why any exposure would be a benefit for his team.
Business Analysts	IT/Business	The Business Analysts suggested that the off the shelf solution would be the most appropriate.

Question 19

Role	Function	Answer
CIO	IT	The CIOs answered this question by referring to the previous question (18). They understand the risks, and they are not prepared to take them.
CFO	Business	On a similar path was the answers from the CFOs. They are not prepared to take any risks.
MDM Directors	IT/Business	The MDM director was the only one that suggested that he would prefer an in-house developed Master Data Management Solution. He supported that each organisation is different and that needs differ from industry to industry. He suggested that he is aware of the risks that an in-house developed solution would involve but he supported that it is more important for him to deliver a solution based on his strategy and not a generic strategy that has been made without taking into consideration the challenges that the organisation that he is part of has. With regards to the lack of
		experience within the IT Development team, he suggested that he would bring in the team external resources to support and supplement the potential gaps that might exist within the development team working together with the CIO and the Head of IT Development.
Head of Development	IT	The Heads of Development answered the previous question by answering that they are not prepared to take any risks.
Business Transformation Program Director	Business	Following the previous answers, all the Business Transformation would go with an off the shelf solution because it is too risky to go with an in-house developed Master Data Management Solution.
IT Project Manager	IT	The IT project manager suggested that they prefer to go with an off the shelf solution because it is less risky than developing one in-house solution
Head of Business Analysis	IT/Business	The Head of Business analysis supported his previous answer where he suggested that the off-the-shelf solution is less risk than the in-house developed Master Data Management solution.

Business Analysts	IT/Business	The Business Analysts suggested that they assume that the risks will be extremely higher developing a solution in house rather than buying an off the shelf solution

Question 20

Role	Function	Answer
CIO	IT	The CIOs were more focused on the infrastructure side of data growth. They said that the could not give an answer based on numbers but they could say that they predict storage growth for each year for multiple systems and of course the data warehouses. The annual OPEX budget for the storage is about 5% of the overall annual IT Budget.
CFO	Business	The interviewer didn't ask this question to the CFOs as he assumed that they were not the appropriate audience.
MDM Directors	IT/Business	The MDM director suggested that annual growth differs per data domain. He suggested that the Product domain
		shows the growth of about 75% year on year. The customer domain shows the growth of about 25% year on year with some existing customers moving to different competition and the Supplier domain shows alteration of about 3% year on year.
Head of Development	IT	The Heads of Development said that in regular times during the calendar year they meet with the infrastructure team and they discuss storage requirements and based on the discussions they decide what additional storage they will need for the next quarter if they need any. The number of records though is not something that they can answer on.
Business Transformation Program Director	Business	The business Transformation Directors suggested that for their projects, the current standard storage prediction that has happened is based on a 3-year plan with a soft evaluation every 3 months and with a detailed evaluation on an anniversary. They couldn't give a specific number of records as an estimation.

IT Project Manager	IT	The Project managers suggested that during the industrialisation phase of a project deliverable they estimate the storage that is required to support the application as well as to operate the application on an annual basis. They couldn't give a range of records that the application would show growth.
Head of Business Analysis	IT/Business	The Head of Business Analysis couldn't answer this question.
Business Analysts	IT/Business	Business analysts could not answer this question.

Question 21

Role	Function	Answer
CIO	IT	Both of the CIOs described the unstructured Data as an aspiration but unfortunately, the organisations that they work are not mature enough for this level of Master Data Management. They both said that they are still struggling with the structured data, so thinking about Mastering the Unstructured data would not be appropriate at this time.
CFO	Business	The CFOs said that they would like to have more insights from the Social Media data but they think that the organisations (that each one of them works) are not ready yet for that level of Master Data Management.
MDM Directors	IT/Business	The MDM director suggested that he would like to have a separate session for the images of the products which is clustered as content and it is managed within the Product Information Management System but apart from that, there is no plan soon to start focusing on Mastering the unstructured data.
Head of Development	IT	Both Heads of Development suggested that there is not any plan for applying any unstructured data in a Master Data Management process.

Business Transformation Program Director	Business	One of the Three Business Transformation Project Directors suggested that there is a plan to analyse Social Media Feeds and categorise them but this is not involving any Master Data Management. The other two Business Transformation Directors suggested that they don't have any plans, for now, to start to utilise in a sophisticated way any social media feeds and there are not any plans in the future to use any Master Data Management on any unstructured data.
IT Project Manager	IT	Both IT project Managers said that there isn't any plan of using unstructured data on any of their projects. And there is no plan to use any unstructured information outside their typical Operational systems.
Head of Business Analysis	IT/Business	The Head of Business Analysis said: "The business needs to optimise the way that they work with the structured data and then think about dealing with unstructured data".
Business Analysts	IT/Business	The Business Analysts said that unstructured data has not been mentioned on any of the processes that have been discussed so far, and that they doubt that there will be any involvement soon.

Question 22

Role	IT	Answer
CIO	IT	The CIOs answered this question in without giving exact numbers, but what they said was that all the internal systems (systems within the Organisation's Network) should integrate with ERP and the ERP should integrate with these systems. Also, they mentioned that in any case that external systems are used, the first point of integration should again be the ERP. Their inspiration is to achieve as much real-time integration as possible, but this is not always possible due to the systems' limitations.
CFO	Business	When the interviewer spoke with the CFOs, he had to change the question due to its technical nature. The question had focused only on the importance of synchronous or asynchronous integration. The CFOs suggested that "as long as the information is accurate, it doesn't matter if the information is one or two days later". All of them would prefer as much real-time information as possible, but they realise that this is difficult with the current technology that the business uses.
MDM Directors	IT/Business	The MDM director suggested that they have "about 25" internal systems that they should integrate. These 25 systems are separated in multiple ERPs, multiple CRMs, multiple WMSs, multiple PIMs, multiple BI platforms and multiple other systems that are used internally and multiple other system integrations that are used externally. Also, due to the nature of the specific Business, multiple EDIs have been set up for a specific type of customers since the organisation is a wholesaler. Some parts of the integrations are as much synchronous as possible but most of the integrations are asynchronous and they run at different times during the day.

Head of Development	IT	Both Heads of Development suggested that the integrations that have been developed cannot be measured but new integrations are being built every day. Depending on the type of integration, the frequency of "the calls" varies but mainly, most of the integrations are asynchronous.
Business Transformation Program Director	Business	The Business Transformation Program Directors said that everything should be going back to ERP and critical information should be sent out to other systems from the ERP. The integration with the Warehouse Management systems should be synchronous and, real-time, however, when someone places an order, the ERP doesn't receive it instantly. The Business Intelligence platforms are being fed information overnight.
IT Project Manager	IT	The IT Project Managers said that they couldn't number all the integrations within the organisation but they could suggest that they were "far too many". The Integrations most of the times is asynchronous and this is due to the capabilities of the systems.
Head of Business Analysis	IT/Business	The Head of Business Analysis suggested that there are multiple integrations between multiple systems. A new integration is designed nearly every day and mainly are asynchronous with most of them though being on multiple schedules within an hour.
Business Analysts	IT/Business	All the business analysts suggested that the integrations are multiple both internal and external with most of them to run on schedules and not on events. What the business analysts' mean by events in real-time synchronous integrations.

Question 24

All the organisations that the interviewees are employed are PCI compliant therefore, all the sensitive information is encrypted and audited in a very regular frequency through the year.

Question 26

Role	Function	Answer
CIO	IT	Both of the CIOs suggested that their organisations at the moment are between the Baseline and the Emerging category by putting their organisation on the step of understanding their Data Environment and measuring the Data Quality.
CFO	Business	All the CFOs suggested that the current stage that they are in is trying to understand the quality of the data within their organisation. They understand and they are aware that there is a problem however, they will need to know the level of this problem.
MDM Directors	IT/Business	The MDM director suggested that he is currently trying to understand the current environment as he is quite fresh to the organisation. Based on this analysis he plans to measure the quality of the data during the analysis of the current environment and then to assess the need of the users who deal with Master Data to go back to the Senior Management to secure their Buy-in.
Head of Development	IT	The Heads of Development suggested that they are currently trying to secure the senior management Buy-in while at the same time trying to measure the quality of the data.
Business Transformation Program Director	Business	What the Business Transformation Program Directors were focusing on is the current quality of the data across the organisation as this is the task that affects their Data Take On process during their Data Transformation Program.
IT Project Manager	IT	The IT project managers suggested that their organisations are in the same positions as the MDM Director. They are measuring the Data Quality while they are trying to understand their current Data Environment. At the same time, they are always trying to make sure that the Senior Manager will support the Master Data Management process and will authorise the budget for the project.

Head of Business Analysis	IT/Business	The Head of Business Analysis suggested that the current state is between Baseline and Emerging as his team of Business Analysts are trying first of all to understand the needs of the business users that they deal with Master Data and at the same time to measure the quality of data across the organisation.
Business Analysts	IT/Business	The Business analysts suggested that the current state of the organisation is the assessment of the needs of the business users on using Master Data and how this need is translated into a process.

Question 28

Role	Function	Answer
CIO	IT	Both CIOs suggested that what they need for their organisation is to implement a formal MDM initiative.
CFO	Business	The CFOs suggested that it is very important to implement a standardised training to the users so they can understand how to comply with a Data Governance policy in their daily duties. It is more important for users to be educated and follow the process.
MDM Directors	IT/Business	The MDM director suggested that the one thing that he needs and he doesn't have at the moment is the support of Senior Management in practice.
Head of Development	IT	The Heads of Development suggested that they should invest more in automation as this would eliminate most of the problems.
Business Transformation Program Director	Business	The Business Transformation Program directors suggested that they would prefer the business to have invested more in automation before the beginning of their Transformation programme as this would save a significant amount of time and money.
IT Project Manager	IT	The IT project managers suggested that what the organisation would need more is to ensure that the senior management would support a project like Master Data Management faster than it takes.

Head of Business Analysis	IT/Business	The head of business analysis suggested that the business should invest more time and effort on training their users follow the principles that a process requires.
Business Analysts	IT/Business	The Business Analysts suggested that Training and automation are the two things that the business should invest in.

Appendix 5.
Interview document

Dear participant,

Thank you in advance for your acceptance to participate in this discussion and thank you for dedicating time to discuss with me these questions. This discussion is recorded, and a transcript of this conversation will be used for my research.

This is the list of the questions that we will discuss.

1. What is your Role

2. What kind of projects have you worked on? What scale?

3. What is the importance of the Data Quality on these projects

4. Are you aware of the classification Master Data

5. What is your definition of Master Data

6. What is your view of Master Data Management and How important is it for your Project?

7. Are you aware of the 3 classifications that a business can be regarding MDM?
    a. Best in Class
    b. Industry Average
    c. Laggard

8. Are you aware of the KPIs with Regards to MDM?
    a. From 1 to 10 (1 low 10 high) what is the mark that you would give to your business with regards to the following?
        i. Formal MDM system in Place
        ii. Automated Capture and Creation of Data
        iii. Cross-functional team both IT and business lines to guide Master Data Management Implementation

9. How would you describe your data strategy with regards to specific business objectives?

10. How would you describe the support from the senior management with regards to the MDM improvement?

11. How would you describe the tools that you implemented to manage and maintain Data Consistency?

12. How would you describe the Cost, the time and the complexity of an MDM implementation and what is the ROI of High-quality data?

13. How much time do you think that an employee of your organisation will spend on searching for data per hour/week
    a. Best In Class is 1.2 hours /week
    b. Industry AVG is 4.4 hours/week
    c. Laggard is 8.2 hours/week

14. Achieving BIC status would save you 355hour/year or 8.9 weeks per year how much would you value the time on searching?

15. Regarding Data sources, what kind of actions are in place with regards to mapping out Data infrastructure to identify high priority data sources and what kind of actions are in place to analyse these resources?
    a. BIC companies spend 48% on mapping out 36% on analysing

16. What kind of difficulties do you face during the implementation? And how can you relate to the most common difficulties?
    a. Data stored in too many silos usually owned by specific division or department or in ERP or CRM
    b. Time
    c. Cost of starting the implementation
    d. Manual Methods of Handling Data
17. How much could you relate to these and how?

18. Have you considered buying an MDM product of the Shelf or building one in-house Solution?
    a. If the product is off the shelf which one and why and if building one in house solution how did you measure the feasibility and why didn't you buy one off the shelf?

19. Are you aware of the risks of developing an in-house MDM Solution?
    a. Usually, it is lack of experience on Understanding MDM, Data governance and access controls
    b. The investigation has shown that 49% of the in-house implementations have shown that there was no improvement on the business or the implementation had negative effects --10% of the businesses that are implementing MDM prefer non -

conventional systems like open source or cloud-based MDM (22% of this 10% is Best in Class)

20. Have you predicted the annual data growth? Usually, it is up to 36% but larger companies can show growth up to 75% or even 150% where would you put yourself?
    a. 40% less than 100000 records
    b. 29% between 100000 and 1000000 records
    c. 31% more than 1000000
    d. (42% of Best in class companies have more than 1000000 records)

21. What about unstructured data formats on MDM? Have you thought about that? and how are you going to facilitate this? Is there a plan to expand unstructured data on your MDM system and how?
    a. (pictures, CAD Files, Facebook or Twitter comments)

22. What kind of Data sources are you planning to use and how many what kind of integration is it planned? Real-time synchronous, or Asynchronous?
    a. BIC usually use 16 unique internal data sources and 7 external while all the other companies use 10 internal and less than 4 external

23. In a scale 1 to 10 what mark would you put on the following? (1 low 10 high)

Importance of Master Data Management to Core business operations	
Buy-in from Senior Management	
Resources/ Budget support for MDM	
Adherence to Master Data policies	
Trust in Master Data	
Trust in Data Systems and Policies	
How likely is to identify Records with significant errors	
How likely is to fix these errors in less than an hour	
How likely is for the information to be gained faster	
How likely is a decision to be made faster	

How likely is to integrate new data sources faster	
How likely is for the employee to reduce the time that he spends on data-centric processes	

24. With Regards to Security and Personal Identifiable Information of Confidential Information, what are the actions that have been considered?
    a. Anything planned with regards to data encryption (what kind of encryption) of any measures to Data Loss protection to prevent unauthorised access and reduce data loss or exposal

25. With a scale 1 to 10 where do you think that you fit on the following table of Best In Class companies checklists

Process	Standardized training for master data system	
	End-user needs for data access and use collected	
Organization	Executive sponsor for MDM	
	The cross-functional MDM team	
	Defined Create, Read, Update and Delete (CRUD) role	
Knowledge	Discovery and identification of all business data	
	Classification and definition of all business Data	
Performance	Measurement tools to track and report data quality	
	ROI for MDM defined and Tracked	
	End-Users time to access master data tracked and measured	

26. There are four business capabilities and enablers with regards to MDM
Where do you think that you stand and why?

Baseline	Organisational Buy-in (Senior Management support, Champion of MDM et)
	Understanding the existing Data environment

Emerging	Assessing the need of end-users that interacts with MDM
	Measuring Data Quality
Early adoptions	Measuring time to Access Information
	Measuring the ROI of MDM
BIC Differentiations	Standardised Training
	Enforcing CRUD (not define but enforcing)

27. Which of the following are you thinking of applying or already apply on your MDM toolbox and from 1 to 10 how much that you use/will use?

Automation	Auto Internal Capture	
	Auto External Capture	
	Auto Indexing – Sorting	
	Auto Cleansing	
Data Management	Data Enrichment	
	Data Governance	
	Data Deduplication	
	Data Cleansing	
	Data Normalisation	
Data Access	Data Access Tools	
	BI for MDM	
	Internal Collaboration Tools	
	External collaboration Tools	
	Mobile Access	

Finally...

28. Which steps do you think that you need to apply to your MDM implementation to improve it? Please select only the ones that you do NOT already have.

Secure Senior Management Support	
Implement a Formal MDM initiative	
Implement Standardised training	
Invest in automation	
Enable Remote Access	

Thank you very much for your time

Panagiotis Lepeniotis

Group SQL Solutions Development Manager

Appendix 6.
Student Ethics Checklist

Research with Human Participants

Question	Yes/No
Does the research involve human participants? This includes surveys, questionnaires, observing behaviour etc.	Yes

Question	Yes/No
1. *Note If YES, then please answer questions 2 to 10* *If NO, please go to Section 3*	
2. Will any of the participants be vulnerable? *Note: Vulnerable' people include children and young people, people with learning disabilities, people who may be limited by age or sickness, etc. See definition on the website*	NO
3. Are drugs, placebos or other substances (e.g. food substances, vitamins) to be administered to the study participants or will the study involve invasive, Intrusive or potentially harmful procedures of any kind?	NO
4. Will tissue samples (including blood) be obtained from participants?	NO
5. Is pain or more than mild discomfort likely to result from the study?	NO
6. Will the study involve prolonged or repetitive testing?	NO
7. Is there any reasonable and foreseeable risk of physical or emotional harm to any of the participants? *Note: Harm may be caused by distressing or intrusive interview questions, uncomfortable procedures involving the participant, invasion of privacy, topics relating to highly personal information, topics relating to illegal activity, etc.*	NO
8. Will anyone be taking part without giving their informed consent?	NO
9. Is it covert research? *Note: 'Covert research' refers to research that is conducted without the knowledge of participants.*	NO
10. Will the research output allow identification of any individual who has not given their express consent to be identified?	NO

Research in Organisations

Question	Yes/No
1. Will the research involve working with/within an organisation (e.g. school, business, charity, museum, government department, international agency, etc.)?	YES

2. If you answered YES to question 1, do you have granted access to conduct the research? *If YES, students please show evidence to your supervisor. PI should retain safely.*	YES
3. If you answered NO to question 2, is it because: A. you have not yet asked B. you have asked and not yet received an answer C. you have asked and been refused access. *Note: You will only be able to start the research when you have been granted access.*	

Research with Products and Artefacts

Question	**Yes/No**
1. Will the research involve working with copyrighted documents, films, broadcasts, photographs, artworks, designs, products, programmes, databases, networks, processes, existing datasets or secure data?	YES
2. If you answered YES to question 1, are the materials you intend to use in the public domain? *Notes: 'In the public domain' does not mean the same thing as 'publicly accessible'.* • *The information which is 'in the public domain' is no longer protected by copyright (i.e. copyright has either expired or been waived) and can be used without permission.* • *The information which is 'publicly accessible' (e.g. TV broadcasts, websites, artworks, newspapers) is available for anyone to consult/view. It is still protected by copyright even if there is no copyright notice. In UK law, copyright protection is automatic and does not require a copyright statement, although it is always good practice to provide one. It is necessary to check the terms and conditions of use to find out exactly how the material may be reused etc.* *If you answered YES to question 1, be aware that you may need to consider other ethics codes. For example, when conducting Internet research, consult the code of the Association of Internet Interviewers; for educational research, consult the Code of Ethics of the British Educational Research Association.*	NO
3. If you answered NO to question 2, do you have explicit permission to use these materials as data? *If YES, please show evidence to your supervisor.*	YES
4. If you answered NO to question 3, is it because: A. you have not yet asked permission B. you have asked and not yet received and answer C. you have asked and been refused access. *Note You will only be able to start the research when you have been granted permission to use the specified material.*	**A/B/C**

Appendix 7.

Practitioner's Recommendations.

The technology is completely different compared to what it was when this research started but the desired outcome is still the same. Organisations need to know how to manage their master data and to achieve that, they need to understand what their master data is. Once these organisations realise that their data is a business asset and not just an IT issue, the improvement within the organisation will be visible immediately, transforming the business from an operationally challenging environment to an operationally enhanced environment. And this operationally enhanced environment learns how to follow the established by data governance rules enabling the organisation to reach operational excellence.

The first recommendation is that organisations need to know their data to understand their business. Without the information that the data will extract, the organisations will struggle to achieve any clear and accurate visibility of their operations, their challenges, their competition, and what their customers need. Data is one of the most important parts of an organisation and as long as the organisations do not treat it as an asset, the organisation will not be able to achieve operational vitality.

Following the above, the first recommendation should be the task to educate the organisation about the way that they use their data as well as to educate the importance of the data to the organisation. This is a challenging task, but it is a necessary task for any data related project.

The second recommendation as defined within this research is the expectations that someone will set at the beginning of any project or programme. The organisation will expect the best possible outcome and process with the minimum effort and the minimum cost, in the shortest period. Unfortunately, these expectations are not realistic for a project that heavily involves both IT and Business. That is why a detailed assessment will have to take place before any commitments. Even if the resources are available and the budget is there, analysis and familiarisation with the environment are necessary to assess correctly what methodology will best fit the needs of the organisation. Based on a different quote during one of the implementations that took place during this research, "A woman needs nine

months to give birth to a baby. Even if you bring 9 women, you still cannot have the baby in one month." (Project Colleague Media Sector 2013). All this analysis, the definitions and the education around the process and the data, will need to happen. It is essential, and it needs to be understood by all the parties that are going to be involved in a project like that.

The Third recommendation that is suggested to anyone that is going to start the implementation of an MDM process is the quote "Don't try to be a hero" (Project Colleague Wholesale Sector 2009). The person who is responsible for the implementation needs to be aware of the limits and capabilities. Each implementer should evaluate the resources, structure a plan, and based on the plan, evaluate the cost, the time, and the feasibility of each task. The act of overpromising sets expectations that are never met and the expectations that are not met are reducing credibility and accountability. The implementer should form a plan based on what is available and present a plan based on what is achievable and not on what the senior management or the sponsors want to hear.

Following on from the third recommendation and moving to the **fourth recommendation**, the implementer should have a strong attitude and do not agree on everything without prior analysis. On the implementation of an MDM process which involves nearly all the functions of the organisation that deals with data within a certain domain that is part of the process, everybody has an opinion. Especially when the implementation is happening in a corporate environment like most of the implementations, any opinion is always strong. That is why the implementer should only agree on what can be achieved and disagree with everything that does not align with the plan. A different point of view should also be raised on everything that falls outside of what it has been agreed on the initial stages of the planning and analysis. By agreeing with every opinion there is a high risk that the objectives and deliverables of the project will be completely different, and usually more, from what it has been budgeted. Also, in a corporate environment, an agreement between all the involved business functions should be reached before any action. If this is not the case, there is a risk for multiple differences between the functions resulting in delaying of the decision making.

The fifth and final recommendation to any MDM implementer is to "follow the process". As it is described in chapter 2, there is a best practice that can be applied in every industry. The

approach and the methodology may vary due to the specific business needs but the approach, the methodology and the implementation can be referenced to a successful implementation. The technology is available to support any implementation style suitable for the organisation's needs. However, the approach is the decision that is made based on the organisation's environment and circumstances and since there is an architecture that can be applied to the organisation's needs, the implementer does not need to face challenges that others have already faced. Based on that, any MDM implementation will have to ensure that the involved parties do not see that as an invasion on their existing process or a suggestion that everything that they used to do up to that point was wrong, but as an improvement.

Appendix 8.
Technical Further work

This research was focused on defining the decisions and the circumstances that these decisions were made during a BTP and how these decisions or circumstances can affect its successful implementation and it did not focus on the actual technology that is used. This research started in 2008. The technology has been substantially improved since then. Cloud computing and cloud processing is something that cannot be characterised as new and revolutionary anymore and most of the organisations are investing in cloud-based services rather than software and especially hardware.

Based on the above statement, there are many services now that offer cloud-based integration between the systems that are used by the organisations. However, MDM is still a process that is kept in-house or infrastructure as a service environment.

Microsoft Azure and Amazon AWS are the strongest vendors of infrastructure as a service as well as database as a service.

Very recently, a very new service got introduced to the public by Microsoft called Azure Data Factory. "*Data Factory is a cloud-based data integration service that orchestrates and automates the movement and transformation of data. You can create data integration solutions using the Data Factory service that can ingest data from various data stores, transform/process the data, and publish the result data to the data stores.*

Data Factory service allows you to create ***data pipelines that move and transform data****, and then run the pipelines on a specified schedule (hourly, daily, weekly, etc.). It also provides rich visualizations to display the lineage and dependencies between your data pipelines and monitor all your data pipelines from a single unified view to easily pinpoint issues and setup monitoring alerts."* (Sharon Lo 2016)

The reason that "data pipelines that move and transform data" is in bold is that this process might be the next generation for the MDM process.

Data factory can work in a hybrid mode which means that can work on both in premises (a server within the organisation's network) and cloud. That means that sources that were out of scope in an MDM process can now be part of the integration and new definitions of master data can arise.

The entire process of the data factory is very similar to the technicalities that Master Data Management can do without the factor of the governance that can be applied. However, a more detailed review of what Data Factory offers is:

1. Data collection from multiple systems cloud and on-premises
2. Ingestion of the data
3. Preparation of the data
4. Transformation and analysis
5. Publishing of the data and
6. Consumption of this data

The equivalent of Master Data Management is:

a. Data Collection from all the systems that have been defined as part of the Master Data Management process.
b. Data insertion into a centralised data repository in which
c. The data is analysed and processed to:
d. Transform the data into the golden record/single version of the truth.
e. Then Master Data Management based on the method of use can publish the golden record back to the source systems for
f. Data consumption by the process and the users.

Comparing the two processes the research can identify that within steps 2, 3, 4, 5 is where the data governance policy can be applied. However, this requires more research.

Data factory's functionality also involves the element of Big Data into the Master Data Management domain since systems of different nature can now be integrated. New master data elements can be introduced into the Master Data Management process including different taxonomies that can be applied to the unstructured data. These taxonomies can be part of the Master Data Management process.

This can be very interesting research, however, the data factory as it stands at the moment, has geographical restrictions as it is only available into the regions of the western United States, eastern United States and North Europe. Also, since the data factory is only a process and not a data repository, it does not store any data in these regions. It only processes these data from one location to another location that the subscriber can specify. As a result, any hesitations with regards to data protection and confidentiality are not valid as the data

can be stored anywhere based on the level of sensitivity and the confidentiality degree that an organisation can define for its data.

In addition to the technological part, further work can be done on the sociology part of any BTP but also on how people react to change and what are the trigger points for these reactions. This could be more of a sociology experimental research with multiple case studies and retrospective interviews.