Journal of Information Systems and Technology Management – Jistem USP
Vol. 19, 2022, e202219017
ISSN online: 1807-1775
DOI: 10.4301/S1807-1775202219017 Original Article

A HEATMAP APPROACH FOR MASTER DATA MANAGEMENT PROGRAMS

AnnMarie Ericsson[1] https://orcid.org/0000-0002-0041-8576
Mikael Berndtsson[2] https://orcid.org/0000-0001-8362-3825

[1]Ammizing Consulting, Skövde, Sweden

[2]University of Skövde, Skövde, Sweden

ABSTRACT

Master data management programs are large by nature since the aim is to provide the entire enterprise with a shared trusted view of the organisation's most critical data assets. In this paper, we present what dimensions and activities a master data management program in a large organisation should consider and how to monitor such a program once it is up and running. Our findings result from an autoethnography study where personal experiences are connected to previous research. A heatmap approach is used to visualize the inherent complexity of a master data management program. Our approach is derived from participating in four different master data management programs in four different global organisations during 2007-2020.

Keywords: data quality; master data; master data management; master data management programs; heatmap

Manuscript first received: 2022-03-09; Manuscript accepted: 2022-10-17

Address for correspondence:
AnnMarie Ericsson, Ammizing Consulting, Skövde, Sweden.
E-mail: ammi@ammizing.se

Mikael Berndtsson, University of Skövde, Skövde, Sweden.
E-mail: mikael.berndtsson@his.se

CC BY

INTRODUCTION

Many organisations are constantly struggling with data quality issues. Poor data quality costs time and money due to inefficiencies in operation, rework, and sub-optimal decisions due to lack of, or wrong information. Additionally, it is a major obstacle for digital transformation. An organisation with poor data quality has problems to adapt to external and internal changes, increased risk of expensive production downtime and may be exposed to legal consequences due to incorrect or missing information.

A common starting point for improving data quality is to focus on improving the quality of data that is of the highest value and is used in the entire organisations. This type of data is named Master Data, and the way organisations manage this type of data is named Master Data Management (MDM).

The vision with master data management is straightforward: create a shared trusted source for storage and distribution of the organisation's most valuable information assets. However, implementing MDM in an organisation is a complex task that affects many systems and processes. Several organisations have fallen into the trap of purchasing an MDM platform and hiring externals to perform the implementation, not realizing that this is a business transformation that potentially will affect the entire business. In most cases, the responsibility for master data will move from IT to business and from local management to central management of master data, which may have a big impact on business processes. Improving master data quality is never a pure IT project, which is also evident in the finding that most of the main barriers to achieving high master data quality are non-technical (Haug, Schlichter, Stentoft Arlbjørn, & Zachariassen, 2013; Haug & Stentoft Arlbjørn, 2011).

The process of setting up and running an MDM program is organisation-specific, so there is no one-size-fits-all solution. Each organisation has different starting points and different flavours of data quality problems. MDM programs are large by nature since the aim is to provide the entire enterprise with a shared trusted view of high-quality data. This can potentially affect all systems, processes, and ongoing projects. An MDM program implies that several or all business areas need to work together with the IT department to create a solid foundation of information during a long period of time.

The organisations we have collaborated with have many different ERP systems, CRM systems, SCM systems, and processes for managing data in different business units. All of which might be the result of merging due to acquisition or lack of data governance. In such a situation, the MDM program is crucial, and a central MDM platform is a prerequisite for managing shared master data. However, an organisation on the other side of the spectrum with a single ERP system and some additional data sources may not need a dedicated master data platform. It may be sufficient with central data governance and the technical support for data management that exists in an ERP system.

Previous research has focused on barriers to master data quality (Haug & Stentoft Arlbjørn, 2011), maturity models (Spruit & Pietzka, 2015; Zúñiga, Cruz, Ibañez, Dominguez, & Moguerza, 2018), and determinants of MDM adoption (Haneem, Kama, Taskin, Pauleen, & Abu Bakar, 2019; Vilminko-Heikkinen & Pekkola, 2019). Although previous research provides valuable recommendations for specific aspects of an MDM program, previous research has not reported how to set up and monitor an MDM program for large organisations. To the best of our knowledge, details of a complete MDM program have not previously been reported in the research literature. In particular, details about dimensions and activities that need to be part of an MDM program are missing. Thus, our research question is: what type of dimensions and activities are needed in an MDM program?

CC BY

This paper presents lessons learned from applying a heatmap approach to MDM programs in large organisations. A heatmap is a two-dimensional approach that can abstract results or the status of a complex phenomenon by constructing a coloured image that gives an immediate understanding of the big picture. Although software designer Cormac Kinney coined the term heatmap in 1991, the origins can be traced back to the early 1870s when shaded matrices were used (Wilkinson & Friendly, 2009). Heatmap in our work was used to structure the work in an MDM program and facilitate a good overview of the status of the MDM program.

In the remainder of this paper, we present a literature review of related work in master data management. Thereafter, we present our research approach and research setting. In the succeeding sections, we present the MDM heatmap, and lessons learned from applying the MDM heatmap. Finally, findings are discussed, and conclusions are presented.

LITERATURE REVIEW

Master data

Master data represents the core business entities of an enterprise along with their associated metadata, attributes, definitions, roles, connections, and taxonomies. Generally, data classified as master data have the following characteristics: it is widely shared and used all over the organisation, it has a long lifespan and a low pace of change, it is part of many transactions, has a lifecycle, and its value for the organisation is high. Examples of master data domains are Things (e.g., Products, Parts, Services), Locations, Roles, Organisation, Party (e.g., Customer, Suppliers, People) (Loshin, 2009). Furthermore, organisations also need to define what attributes and metadata attributes are classified as master data and identify how they are connected. It is, for example, not economically efficient to *master* all attributes an organisation holds about a product. Everything is not important for everyone, so each organisation needs to define what attributes shall be treated as master data for them. Some common attributes to master for products are identifiers, weight, and size measures.

For organisations where the speed of updates of master data is important, Karia, Sundararajan, and Raghavan (2021) have proposed an approach based on distributed ledger technology and recommendations on selecting suitable business processes that would benefit from such an approach.

Master data quality barriers

In a literature review, Haug and Stentoft Arlbjørn (2011) derived five general barriers to master data quality: i) lack of delegation of responsibilities for maintenance of master data, ii) lack of rewards for ensuring valid master data, iii) lack of master data control routines, iv) lack of employee competencies and v) lack of user-friendliness of the software that is used to manage master data. In a follow-up investigation (Haug et al., 2013) targeting the main barriers for Danish manufacturing companies to achieving high master data quality, 12 barriers were identified:

- Missing placement of responsibilities for specific types of master data
- Lack of clarity of roles in relation to data creation, use, and maintenance
- Inefficient organisational procedures

CC BY

- Lack of management focus in relation to data quality
- Lack of data quality measurements
- Lack of rewards/reprimands in relation to data quality
- Lack of training and education of data users
- Lack of written data quality policies and procedures
- Lack of emphasis on the importance of data quality by managers
- Lack of IT systems for data management
- Lack of possibilities for input (e.g. fields) in existing IT systems
- Poor usability of IT systems

In both investigations (Haug et al., 2013; Haug & Stentoft Arlbjørn, 2011), the non-technical barriers are in the majority. The authors experienced most of the barriers described above except "Lack of management focus in relation to data quality" and "Lack of emphasis on the importance of data quality by managers" if upper management is considered. This might be because these organisations were running a large-scale MDM program, which is impossible without strong support from upper management. However, even with strong support from upper management, middle management may be reluctant to support the MDM program if it provides no benefits or drawbacks for the managers' unit. This barrier was experienced from middle management. Additional barriers found are lack of trust in central data management and organizational politics.

Master data management

Master Data Management is a common starting point for improving data quality and is an approach for creating a trusted source for storage and distribution of an organisation's core business entities, i.e., master data. Sample definitions of master data management are presented in Table 1.

Most definitions in Table 1 emphasize that MDM is a process for managing and improving master data quality in an organisation. A process that requires business units and IT to collaborate and where the purpose of the technical platform is to support the business process for managing master data.

In addition to using MDM within an organisation, it can also be applied to open data sources (Cadena-Vela, Mazón, & Fuster-Guilló, 2020) and smart infrastructure development (Yang, Wen, Aziz, & Luhach, 2021).

MDM maturity models

MDM maturity models are an approach for assessing an organisation's overall status of ability to manage master data. The maturity assessment can be an initial step to improve the master data quality or as an assessment after activities have been done to improve the master data quality. To the best of our knowledge, only two MDM maturity models have been suggested in the research literature: MD3M (Spruit & Pietzka, 2015), and a maturity model for microfinance (Zúñiga et al., 2018). In contrast, there are a plethora of MDM maturity models suggested by various vendors.

CC BY

Table 1. Sample definitions.

Reference	Definition
Berson & Dubov, 2007	Master Data Management (MDM) is the framework of processes and technologies aimed at creating and maintaining an authoritative, reliable, sustainable, accurate and secure data environment that represents a 'single version of truth,' an accepted system of record used both intra- and interenterprise across a diverse set of application systems, lines of business, and user communities.
Haneem et al., 2019	MDM is not just about the technology. It is about management of shared master data to reduce redundancy and ensure better data quality through standardized definitions and use of data values with a combination of process, governance, and technology.
Smith & McKeen, 2008	Master data management (MDM) is an application-independent process which describes, owns, and manages core business data entities. It ensures the consistency and accuracy of these data by providing a single set of guidelines for their management and thereby creates a common view of key company data, which may or may not be held in a common data source.
Spruit & Pietzka, 2015	Master Data Management (MDM), which we define as "the management of the consistent and uniform subset of business entities that describe the core activities of an enterprise."
Vilminko-Heikkinen & Pekkola, 2019	Master data management (MDM) is a data management practice aiming to ensure the quality of data by managing the organization's core data, referred to as "master data.
White, Newman, Logan, & Radcliffe, 2006	MDM is a workflow-driven process in which business units and IT collaborate to harmonize, cleanse, publish and protect common information assets that must be shared across the enterprise. MDM ensures the consistency, accuracy, stewardship, and accountability for the core information of the enterprise.

Spruit and Pietzka (2015) have proposed a maturity model MD3M that includes five maturity levels, five dimensions (focus areas), and 13 criteria (sub-topics). The MD3M model has been applied and validated in several contexts, e.g., energy sector (Spruit & Pietzka, 2015), public hospital (Rahman et al., 2019), infrastructure networks for banks (Iqbal et al., 2019), presidential advisory council (Ko, Adywiratama, & Hidayanto, 2021), and in Statistics Indonesia (Krismawati, Ruldeviyani, & Rusli, 2019).

Zúñiga et al. (2018) have suggested an MDM maturity model that includes five maturity levels, six dimensions, and 14 evaluation criteria. The maturity model has been validated in a Peruvian microfinance institution.

An overview of the levels, dimensions, and criteria for the MDM maturity models are presented in Table 2.

MDM programs

Implementing an organisation-wide MDM program leads to different benefits in different business units. Some units will experience efficiency gains since people do not have to copy-paste information between Excel files or spend time searching for missing or erroneous data. Other units will get solid and trusted data on which they can implement advanced digital solutions, optimize warehouse inventory, automate processes, avoid legal lawsuits, or have a smoother implementation of new ERP systems.

CC BY

Table 2. MDM maturity models suggested in the research literature

	Levels	Dimensions	Criteria
MD3M (Spruit & Pietzka, 2015)	1. Initial 2. Repeatable 3. Defined process 4. Managed and measurable 5. Optimized	1. Data model 2. Data quality 3. Usage & ownership 4. Data protection 5. Maintenance	1. Definition of master data 2. Master data model 3. Data landscape 4. Assessment of data quality 5. Impact on business 6. Awareness of quality gaps 7. Improvement 8. Data usage 9. Data ownership 10. Data access 11. Data protection 12. Storage 13. Data lifecycle
Microfinance (Zúñiga et al., 2018)	1. Initial 2. Managed 3. Defined 4. Quantitatively managed 5. Optimized	1. Policies 2. Data governance 3. Data models 4. Data integration 5. Data quality 6. Monitoring	1. Policies, procedures and methods 2. Roles and responsibilities 3. Use and ownership 4. Master data identification 5. Master data dictionary 6. Landscape data 7. Master data integration 8. Master data alignment 9. Data integration designs 10. Master data life cycle 11. Identification of quality barriers 12. Quality evaluation 13. Improvements in quality 14. Monitoring and Reporting

Setting up and running an MDM program is a big task with an inherent complexity that comes with a program that potentially affects all areas of an organisation. The exact setup of processes and technology for achieving a successful MDM implementation is organisation-specific, so there is no one-size-fits-all solution. Many organisations fail to fully implement MDM since they fail to recognize that many components, functions, and subject areas are required to make MDM work correctly or are created due to an integrated set of disciplines and shared data (Allen & Cervo, 2015).

Several aspects can affect successful MDM adoption. Haneem et al. (2019) investigated determinants of MDM adoption in the context of Malaysian local governments. Their findings were that the following determinants had significant effects on the MDM adoption: complexity, data quality, data governance, top management support, technological competence, and citizen demand. Determinants that did not have significant effects to MDM adoption were: relative advantage, data security, and government policy.

Similarly, Vilminko-Heikkinen and Pekkola (2017) observed two MDM projects in the public sector and identified 15 challenges. Out of the 15 challenges, eight of them were categorized as MDM specific: concepts related to MDM, MDM concept owner, unified terms and concepts, level of granularity for defining data sets, roles and responsibilities, data ownership, legislation-driven challenges, and mutual understanding of master data domains. A follow-up study was done on

CC BY

how roles, responsibilities, and ownership in the organisation changed during MDM development (Vilminko-Heikkinen & Pekkola, 2019).

In the research literature, we have not come across descriptions of complete MDM programs (with dimensions and activities) for large organisations.

MDM implementation styles

Gartner (Radcliffe, 2004) suggested four different approaches of implementing MDM, and they are now considered as de facto standards in the industry:

- The Coexistence approach consolidates master data from different sources in an MDM platform. A workflow-supported process can be set up to manually validate data before distribution to downstream systems and manage records where the technical algorithms cannot automatically create a golden record.
- The Centralized approach moves all management of master data to the MDM platform, which provides a single implicit source of truth from which all master data is managed and distributed. This approach provides high control of the data since it reduces the risk of introducing errors by providing workflow support with approvals. However, it implies that all management of the master data must be performed in the MDM platform, which may have a big impact on the organisation's business processes.
- In the Registry implementation style, only the key identifiers of the master data entities are stored in the MDM platform. This is the least intrusive implementation of MDM and maybe the only solution if, for example, legacy reasons prevent data from being moved. In the Reference implementation style, master data is gathered and consolidated by the MDM platform from source systems only on request.
- The Consolidating approach data is consolidated from various sources and cleaned and harmonized in an MDM platform. It creates a centralized repository of master data for reporting and analytics, but operational systems are not using the information for their processing. In this approach, data is merged and harmonized, but any errors discovered in the data need to be changed in the source system.
- Centralized and Coexisting implementation styles imply that master data is distributed from a trusted single source, and data is not distributed to downstream systems unless it has been verified for data correctness. Consistency is achieved since only the MDM platform can distribute master data changes to downstream systems. Central and Coexisting will provide master data with high quality to operational systems and have the potential to increase efficiency in operation and compliance to legal requirements all over the enterprise. However, they will also affect many downstream systems and how people work with data, so the implementation may be complex. Registry and Consolidated still allow master data to be inconsistent and even wrong in different systems, but they are less intrusive and easier to implement. Hence, the Registry and Consolidated styles are suitable if master data quality is only required for analytics and reporting, but they can also serve as a first strategic step towards implementing Coexisting or Centralized.

CC BY

Heatmaps

A heatmap is a visualization technique that uses color to highlight individual values that describe the relation between two dimensions. For example, a simple table with rows and columns where some cells are colored (based upon their value) is a heatmap. According to Wilkinson and Friendly (2009), the origins of heatmaps can be traced back to the early 1870s where a shaded matrix was used to highlight characteristics of maps of Paris. Heatmaps are widely used in natural science (Wilkinson and Friendly 2009) but have also been applied to domains such as software engineering (Ishizuka, Washizaki, Fukazawa, Saito, & Ouji, 2019). In particular, clustered heatmaps have become popular since they can group similar rows (columns) near each other and show hierarchical cluster trees along the margins (Wilkinson & Friendly, 2009).

A heatmap can support the decision process in projects where the information is large and hard to monitor. In this paper, we use a simple heat map to follow the status of a complex Master Data Management program. Even though the complexity of an MDM program is not in the range of the complexity of problems appearing in some natural science problems, the number of people, business units, and processes that need to synchronize in such programs make it hard for everybody to share the same view of the program status. Simple heatmaps are intuitive and easy to understand and require little or no explanation, making them ideal for showing progress in, for example, a steering group meeting where time is often limited.

RESEARCH SETTING

To investigate the academic research on master data management program, we used the following databases: ACM Digital Library, Emerald, IEEE Xplore, SAGE journals, Science direct, Scopus, and Taylor & Francis, with the search strings “master data management program”, “master data management project”, “MDM program”, and “MDM project” to investigate if any of these phrases was mentioned in the title or abstract. Four articles were discovered (Cleven & Wortmann, 2010; Vilminko-Heikkinen & Pekkola, 2013; Vilminko-Heikkinen & Pekkola, 2019; Zúñiga et al., 2018), but none of them presented details of a master data management program. This contrasts with many MDM programs offered by consultancy firms and organisations that initiate MDM programs to improve their big data initiatives. Our conclusion is that master data management programs is an unexplored subject in academic research.

This paper is the result of autoethnography research (Tony, Stacy Holman, & Carolyn, 2015), where the authors have reflected on the experience and results gathered from four master data programs in four Swedish global organisations. A similar approach was taken by Vilminko-Heikkinen and Pekkola (2017, 2019), that used an ethnographic research approach to investigate master data management within the public sector.

An autoethnography approach was chosen as it allowed us to synthesise 13 years of personal experiences acting as an architect or project manager in MDM-projects with academic research results. According to Myers (1999), “[t]he main difference between case study research and ethnographic research is the extent to which the researcher immerses himself or herself in the life of the social group under study”. Given the length of the personal involvement and rich data collected by personal observation, a case study with interviews was not considered optimal.

CC BY

Research data (results of completed MDM-projects, project documentation, and personal observations) was collected 2007-2020 during participation in four different MDM programs in four different global organisations with revenue between 1,5 to 40 billion EUR:

- The first project was an implementation of a multi-domain MDM system, starting with supplier and product domain. The scope of the program was to migrate from centrally managed legacy PIM and MDM platforms into a single new multi-domain MDM platform and set up an information governance forum spanning over the entire information domain of the enterprise. The Co-existing implementation style was used in this project. This company has several production sites, sales offices, and franchisees, and their customers are both other businesses and consumers.
- The second project was an implementation of a central product and supplier MDM platform, where no legacy MDM platform or processes previously existed. This project used the Central implementation style. The scope of the project was the entire company. This company has several production sites and sells their finished products both through franchises and to contractors.
- The third project did not include any implementation of a platform; it was an MDM assessment on how to manage master data in legacy systems without implementing a new platform or by implementing a platform with Consolidating implementation style. This project focused on a single production site in a global car manufacturer.
- The fourth project was a major project focusing on changing the central MDM platform, product structure, and processes for managing product information. This project used the Co-existing implementation style. This company is the franchisor of many franchises that should reuse the master data from the franchisor.

The 1st author acted as responsible MDM and/or Information architect in all cases, the MDM heatmap was developed and refined during these projects. Although quite different in project setup, project management, and performance, all four projects aimed to improve data quality issues of supplier and product information. Compared to managing product information, supplier information is less complex, so in most cases, that was the first domain to implement. The aim of the first version of the heatmap was to educate and explain to management how master data must be managed and why it is important to have a holistic view of data quality problems and not just focus on technical issues. In the fourth project, the MDM heatmap was refined and used for structuring the early phases of the project in different workstreams and explaining status of the different parts of the program to management. Although useful in the smaller project, the heatmap was even more useful in the largest project since it is often hard to keep track of and get a simple overview of a project that is divided into several workstreams.

THE MDM HEATMAP

The MDM heatmap provides a simple, yet powerful visualization of the inherent complexity of an MDM program. The intention of the MDM heatmap is twofold: i) guide organisations in planning

CC BY

and performing an MDM program, and ii) visualize the status of the progress once the MDM program is up and running.

The MDM heatmap is composed of eight dimensions, where each dimension can be viewed as a grouping of activities that need specific attention when they are done by the organisation. As shown in Table 3, dimensions can be divided into three categories with respect to main responsibility.

Table 3. Dimension categories.

Business	**Joint business & IT**	**IT**
Vision & business alignment	Information	Integration & Synchronization
Planning & change management	Application	Migration & Roll-out
Governance		Technology

A high-level view of an MDM heatmap is shown in Table 4, where the colour of the cell indicates the status of the activity: no colour (activity not started), green (activity on track), yellow (activity behind schedule or minor issue), and red (major issues with activity). The heatmap was created from the experiences gained in our first two projects and successfully used in project three and four for monitoring and reporting project status.

Table 4. MDM heatmap.

	Activities				
Vision & business alignment	Vision	Scope	Strategic drivers & business case	Business alignment	Business capabilities
Planning & change management	Roadmap	Business process	Stakeholder mapping	Change management	Business vocabulary
Governance	Organisation & decision framework	Definitions, policies, principles & guidelines	Business rules	Roles & responsibilities	Data ownership and data stewardship
	Tool support	Performance and sustainability	KPI metrics	Resource allocation	Risk Management
Information	Master Data Definition	Data Models	Hierarchy & classifications	Data Lineage	Entity resolution
	Data Quality	Metadata management	Data Security	Availability, Scalability & Performance	Process automation
Application	As-Is & To-be System topology	Application security	Platform responsibilities	MDM Implementation style	External Portal
	Multidomain MDM	Role of Analytics & BI	Requirements on MDM platform	Workflows & approval flows	MDM specific application capabilities
Integration & Synchronization	Integration pattern	Integration strategy	Data distribution	Message format & schema mapping	Integration platform
Migration & Roll-out	As-Is To-be data mapping	Load and cleanse/ match and merge	Data migration plan	Roadmap / implementation plan	Data cleansing
Technology	Technical enablement	Non-functional requirements	Platform requirements	Technical limitations	Technical components

CC BY

As most challenges with implementing an MDM program are non-technical, we chose to focus on the dimensions that have a strong business focus in subsequent sections.

Vision and business alignment

The purpose of the dimension *vision and business alignment* is to anchor the MDM program as business driven. Failing to treat MDM as a business program is the most common reason for an unsuccessful implementation since the greatest challenges of MDM are not technical but organisational (Haug et al., 2013; Haug & Stentoft Arlbjørn, 2011).

Ensuring support from upper management as well as connecting and defining a joint **vision** about how MDM should support the overall business strategy are essential keys to success. When the vision is unclear or not communicated and supported by top management, people tend to steer the work they do for MDM to fit their own, often siloed, targets. This will lead to sub-optimization, debates, and conflicts regarding everything from "what master data is" to "how to set up the MDM governance".

Uncertainties in the **scope** of the MDM program must be resolved as soon as possible. The scope defines what type of data is covered, for example, if the target is a multi-domain MDM or only concerns one data domain if the data governance concerns the entire organisation or just a section, responsibilities between the technical MDM solution and other data platforms, how to incorporate capabilities of legacy systems. These big questions will slow down the MDM program if the scope and interfaces between these areas are not agreed upon upfront.

Although most people intuitively understand the importance of high-quality data, the return of investment of an actual implementation of MDM in a specific organisation may be hard to specify. Reasons for this are that most benefits of implementing MDM are indirect since MDM is the foundation for other activities. Hence it is important to consider supporting **strategic drivers & business case**. Delivering MDM to support **strategic drivers & business case** will provide a stable long-term solution that can be reused by other initiatives. This must be clearly explained in the business case for MDM.

The expected benefits must be communicated and directed to key stakeholders in different business areas to ensure that everybody understands "what's in it for me". Working with **business alignment** is to share the MDM way forward, align to what MDM is today, educate stakeholders, explain why the organisation needs to change, and articulate the expected benefits for each part of the organisation. Hence, it is necessary to direct the expected benefits of MDM in different business areas, or middle management will soon try to prioritize other tasks that bring more immediate value to their area of responsibility.

If the organisation has defined a set of **business capabilities** mapped to business targets and processes supporting them, then business capabilities that are affected by the MDM program need to be identified. In this paper, business capabilities represent *what* the business does, not how or by whom. Properly defined business capabilities should be more stable over time than, for example, business processes, since *what* a business does changes more rarely than *how* it is done.

In our experience, an MDM program without clearly identified benefits in all business units will encounter difficulties when resources from, e.g., local units need to change their way of working. Previously they had control over their local master data, and now they must change to comply with a central process in a central system. Furthermore, they can no longer easily change their master

CC BY

data and may also have to reuse data created by others. Even if everybody understands the central benefits of an MDM program, it is not always obvious what the benefits are for the local units. The shift from local management to central management often causes organisational resistance. This shift is illustrated in Figure 1 where ownership is represented by grey boxes.

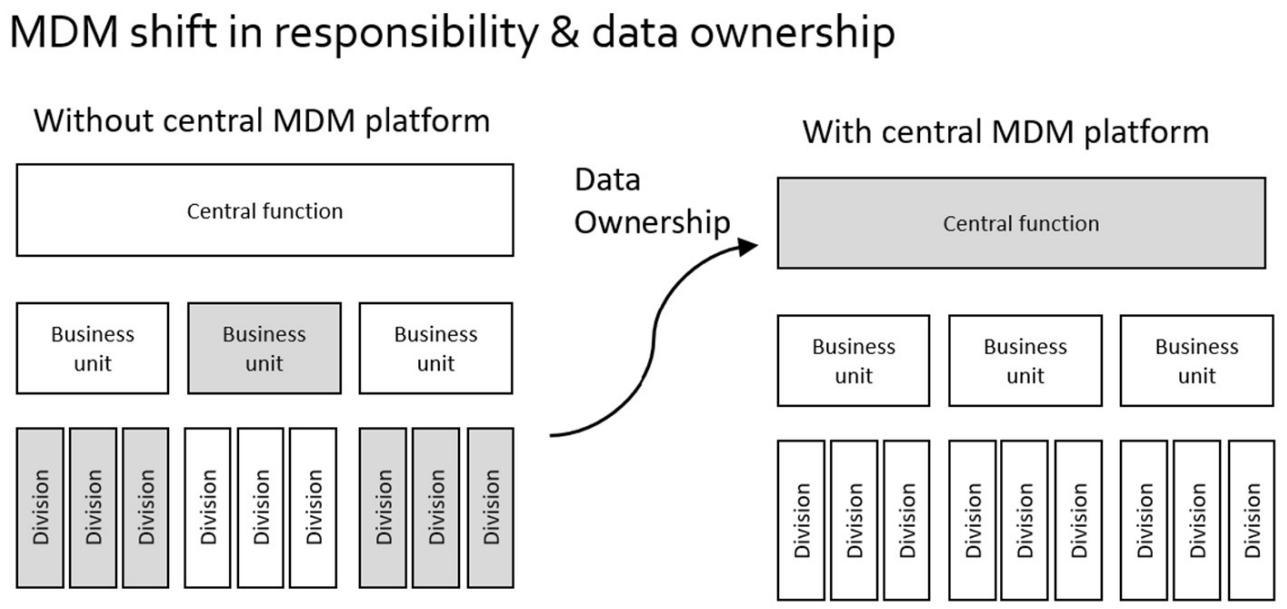

Figure 1. MDM shift in responsibility & ownership.

Planning & change management

The purpose of the dimension *planning & change management* is to develop an overall roadmap and identify main stakeholders and business processes that are affected by the MDM program.

A strategic MDM **roadmap** needs to be developed that will outline how to sequence the work from the current state to the desired vision. Like any roadmap development, the MDM roadmap must be realistic, state the value of each phase, and have milestones. It must be possible to measure the progress of each milestone and each milestone must have assessment criteria, i.e., what needs to be completed for the milestone to be considered complete. Ideally, each milestone should also provide a concrete business value. Defining an MDM roadmap will enable the decomposition of the vision in sub-projects, ensuring that all sub-projects fulfil a part of the shared vision.

Current **business processes** associated with master data management must be documented together with the information exchange between activities in the process flow, which roles that are involved in the processes and the responsibilities for each role. The current **processes**, **roles and responsibilities** need to be mapped to the future processes, roles and responsibilities and a plan must be defined for what efforts in education, alignment, and information that needs to be included for people who are affected by the changes of going from as-is to to-be processes. MDM will change the way people operate, since implementing MDM will typically move responsibilities and management of data from a scattered solution where different roles managed different versions of the same data object, to a central management of a single shared version as illustrated in Figure 1 (assuming Central or Co-existing implementation styles).

CC BY

Knowledge of your stakeholders and their expectations of the MDM program is highly important for a successful implementation. Conducting a **stakeholder mapping** is a prerequisite for business alignment and includes creating a structured overview of the stakeholders and their expectations. This will also enable tailored communication to different groups of stakeholders in order to ensure that stakeholders' expectations are in line with the MDM vision.

Change is a fact in all projects, so include **change management** in the plan from the start. Education and training are paramount for users to feel confident with the change and understand what is happening and why. Consider using an agile framework due to their ability to manage change and strive to delegate decisions as much as possible. Nevertheless, when teams do not agree, there must be an escalation path and a forum for raising conflicting requirements or disagreements in priorities. Solving cross-functional issues is one of the responsibilities of the data governance structure, so the change management escalation path should ideally end in one of the governance forums.

In an enterprise-wide program, it is crucial to establish a common **business vocabulary** and a common language to minimize misunderstandings and misinterpretations between stakeholders. Shared vocabulary and definitions are crucial in MDM programs but may be hard to achieve over an enterprise that has used local variations of definitions and language over a long period of time. It may even be difficult to agree on how to define something as foundational as a customer.

In our experience, starting with a companywide glossary for common concepts will reduce the number of misunderstandings in communication going forward. A clearly defined roadmap with milestones that states business value will reduce the risk of upper management losing patience in the program even if it spans over a long period of time, since it will continuously deliver business value.

Governance

The purpose of the dimension *governance* is to establish principles for data governance.

Large organisations need a multi-levelled **organisation & decision framework** and associated escalation paths. For example, a *strategic steering committee* populated by executives from all different business units as well as IT. This is the highest authority of data governance, and this is where the funding of cross-functional initiatives should be approved and controlled. In contrast, a *tactical data governance committee* focuses on breaking organisational data silos, prioritising cross-functional data issues, and developing and monitoring the data governance policies and principles. A multi-level governance organisation can be useful as a target or vision; however, in the beginning of an MDM program this is often too far from current reality to make sense. Most decisions about data and data governance will be taken by people working daily with data issues. Enforcing consistent use of data over an entire organisation requires a structure and escalation path that keeps data-related issues together, defines policies, and governs that the policies are followed. This is where the enterprise-wide governance organisation is needed. In the beginning, data governance may just be a forum populated by people from different part of the organisation that is equipped with the mandate to dictate, for example, what data stewards must do and what IT shall deliver in terms of technical support.

The backbones of data governance are the **definitions, policies, principles & guidelines** that the organisation wants to impose on data to increase efficiency in data usage, data security, or other characteristics concerning data. The focus of data governance may be different in different

CC BY

organisations. For enterprises covering several competing brands, sharing information over division borders may be a big issue, while a company with a single brand may have problems regarding conflicting laws and regulations to fulfil for the same product in different markets. The first step to successful data governance is to have a common view in the organisation of what data governance means, what should be governed, by whom, what value it will bring to the business, what policies and rules data shall comply to, etc. Once these questions are answered, it becomes easier to motivate people to take on roles in a data governance organisation.

Frequently, information is managed by a set of **business rules**, e.g., defining a maximum set of different products for a certain kind of store or set a priority to customers that match certain characteristics. Business rules need to be visible, understood, and possible to manage and control, and not buried in code or implemented as stored procedures in a database. Identifying and defining the complete set of business rules for a large enterprise is a major task, but it is necessary to ensure correct input and management of information. What is correct in one sub-organisation may be wrong in another due to different business rules. Data compliance to business rules is one aspect of an organisation's data quality.

The data governance organisation will be populated by different **roles & responsibilities** for different aspects of data. It is advisable to appoint information owner for each data domain as early as possible. This is a role with a mandate to decide cross-functional issues about data domains which will facilitate progress when data issues occur. A clear set of data management roles and responsibilities for each role can be difficult to implement in an immature organisation and may require substantial change management efforts to succeed.

The **data owner** has the mandate to decide cross functional issues about the data domain he or she owns. A data owner accepted as an authority by all business units may reveal a lot of the political hassles that may occur along the way and help the program to run efficiently. The **data steward** is a subject matter expert in the data domain to be implemented, e.g., Customer, Product or Supplier. Typical tasks of a data steward are to continuously analyse data to improve quality and ensure compliance to the business rules surrounding the data. Identifying existing data stewards (or similar roles) in the organisation, their current responsibilities and how they work with master data is a first step towards implementing data stewards at the tactical level of master data governance.

The level of **tool support** needed for master data management need to be defined. A high level of technical support may automate a large amount of manual work and facilitate on-boarding of new employees. However, it may also introduce a rigid process leaving little freedom for experienced employees to be effective. It may also decrease quality in some cases where users are forced to type in bad information since data is missing but the field is mandatory.

Means of measuring **program performance and ensure sustainability** of the MDM work needs to be in place. This to ensure progress and momentum even after the MDM program has finished. Hence, while the program is running, budget and resources are allocated, however, it must be ensured that a sustainable organisation exists with budget and resources for keeping momentum after the MDM program has ended, otherwise it will not take long until data quality are degrading again.

To ensure progress in data quality it is important to define measurable **KPI metrics**. Identify what KPI metrics that will drive the work of data quality forward, set targets for each role and

CC BY

measure progress periodically. A KPI measure may, for example, be data coverage of a certain product assortment, number of reported data issues for a customer segment or compliance to business rules.

As in any program, **resource allocation** is necessary. Start with identifying the required expertise needed and ensure allocation from the correct business roles. Data governance issues are close to the core of the business, and it is paramount that people working with data governance have good knowledge about business and a good understanding about the business information flow. People working in an MDM program must have a good knowledge about how the business works end to end, and what information that is used where.

Risk management is important in an MDM program since deploying an enterprise-wide MDM program potentially affects all business processes and all systems in the entire enterprise. Information previously managed differently by different business areas should now be consolidated and everybody must follow the same process for master data management. The work of consolidating all the data and technology can be huge but may still be a small thing compared to sorting out the political and organisational differences that may arise in this exercise. Hence, the risk assessment for MDM must both consider technical issues, data security and risks imposed by changing organisational processes and responsibilities.

In our experience, data governance is the most important and most difficult task of an MDM program. In organisations with few or no strong central governance forums it may be a good idea to let the data governance start with a common concrete task, such as defining a common product classification. This immediately raises questions about who has the authority to decide this and how shall the result be communicated and enforced.

LESSONS LEARNED FROM RUNNING MDM HEATMAP PROGRAMS

Over the years, we have gathered practical experience from several large MDM programs in different companies. We have encountered the barriers as reported by (Haug et al., 2013)Vilminko-Heikkinen and Pekkola, 2017) as well as barriers stemming from more political and organisational issues. Below we describe our lessons learn for mitigating some of the most common barriers.

Lack of management support and business resistance are commonly reported barriers (Haug et al., 2013)Vilminko-Heikkinen and Pekkola, 2017). Our experience is that these barriers can be mitigated by agreeing on why MDM is needed and delivering business value early. A clear vision and roadmap for MDM must be developed and communicated. MDM means different things for different people and is implemented in different ways for different purposes, so it must be clearly defined what the expected outcome of MDM is, and why the organisation invests in it. Delivering actual business value early in the MDM project will reduce the risk of sponsors losing patience in the project. Hence, include early delivery of business value in the project plan.

MDM is not yet another IT project. Vilminko-Heikkinen and Pekkola (2017) observed two MDM projects and "[t]he MDM projects were regarded as IT projects.". This in turn led to resistance from business people, and in particular, business process owners. Although the MDM heatmap does not impose any strict sequential order of the activities, a strong recommendation is to start with the more business-oriented dimensions and activities before approaching the more technical dimensions and activities. The reasons are that the more business-oriented dimensions and activities are frequently

CC BY

prerequisites for the more technical dimensions and activities. Starting with the dimension *vision and business alignment* will also kill, or at least shorten, some of the discussions arising during project execution when the concepts, definitions and targets of master data shall be defined and implemented. Finally, including and guiding business users is one way to introduce users to work with improving data quality.

Monitor activity status patterns. A successful MDM program will have an activity status pattern where the business-related activities are slightly before the IT related activities. In the first phases, only business activities need to be ongoing. IT will preferably start with as-is analysis and investigating data lineage on current master data if that is not already properly documented. In contrast, if the starting point of the MDM program is IT-related, for example, that the purchasing of a new MDM platform is done before any business-related work has been done, then the risk of failure for the program is sever. If the IT related parts of the heatmap is before the business-related parts, then the risk of wasting time and money on implementing a solution that solves the wrong problem is high. It is also a high risk that the program is regarded as an IT program that IT can solve without much business involvement. An organisation's information is owned by the business so business experts must be in the driving seat of these programs to ensure success.

The activity patterns going on in an MDM program are not easily detected without a heatmap, since activities occur in different departments and different parts of the organisation. Using the MDM heatmap for tracking status is a good indicator to see if the program activities are in good balance.

DISCUSSION

Although the MDM heatmap shares several similarities with MDM maturity models (Spruit & Pietzka, 2015; Zúñiga et al., 2018), MDM maturity models have a different purpose than the MDM heatmap. While the MDM maturity models aim to assess the maturity of the master data in an enterprise, the MDM heatmap intends to support the planning of a concrete MDM program as well as provide a visualization of the changing status during execution of such a program. Consequently, the MDM heatmap includes a broader range of activities than the maturity models since an MDM program needs to support the activities enabling an organisation's transformation from one level of MDM maturity to another. The MDM heatmap is based on experience combined with research but may still need to be tailored for specific needs in specific projects.

We agree with previous research (Haug et al., 2013; Haug & Stentoft Arlbjørn, 2011) that the non-technical barriers to achieving high master data quality are in the majority. All of the non-technical barriers identified are issues that must be dealt with, and we have experienced them in practice. However, the most problematic issues in our experience are political. Implementing a modern MDM platform with new MDM processes may, in some cases, mean that people who have built their careers on the know-how of master data lose their power. Furthermore, data can be democratized and made available to anyone without going through a certain business unit. Managers that understand the power of data also understand the power of owning and controlling the platform and processes around that data. In a matrix organisation or an organisation consisting of competing brands, this is not an easy landscape to navigate in, and it puts even more emphasis on vision and alignment. Strong top leadership directions may not be enough if the company's culture is built on entrepreneurship where every sub-part of the business optimizes towards its agenda. Commitment is required from all included areas and an understanding and willingness to drop the local optimization and work for the

CC BY

greater good of the entire enterprise when it comes to data management.

When it comes to the concrete issues identified by (Spruit & Pietzka, 2015; Zúñiga et al., 2018) and Vilminko-Heikkinen and Pekkola (2017), they are all included in the MDM heatmap in a structured way, some as separate cells and some as activities in a cell. Failing with any of the cells in the MDM heatmap will be an issue to the MDM program and although we have emphasized the business parts in this paper, the program will not succeed without the IT related cells being completed in a good way. Hence, using the MDM heatmap will help program managers to follow up coverage of all aspects of the program and avoid or mitigate the issues.

Combining our experience with the findings from (Spruit & Pietzka, 2015; Zúñiga et al., 2018) and Vilminko-Heikkinen and Pekkola (2017), we would like to specifically emphasize the importance of appointing a committed **data owner** that is accepted as an authority by all business units, and that has the ability to drive MDM related questions for his or her data domain across the enterprise. This role may be the difference between success and failure for the entire program since (s)he is an important player in sorting out the political issues that may be impossible for the MDM program to solve.

CONCLUSIONS AND FUTURE WORK

In this article, we presented a heatmap approach for MDM programs in large organisations. The approach is based on our experience from setting up and running four MDM programs in global organisations.

Previous research has focused on barriers to master data quality, maturity models, and determinants of MDM adoption. The significance of our research is that it extends previous research with details about what dimensions and activities an MDM program in a large organisation should consider and how to monitor the MDM program once it is up and running. We have not come across similar descriptions of complete MDM programs (with dimensions and activities) in the research literature.

Future work will include a deeper analysis of each cell in the MDM heatmap and case studies using the MDM heatmap in projects with different implementation styles. Most of the discussions in this paper assume a Central or Co-existing implementation style since they are most intrusive for business processes and legacy systems. However, most of the business part of the MDM heatmap needs to be performed even before the correct implementation style can be chosen.

REFERENCES

Allen, M., & Cervo, D. (2015). *Multi-Domain Master Data Management : Advanced MDM and Data Governance in Practice* (Vol. First edition). Waltham, Massachusetts: Morgan Kaufmann.

Berson, Alex, & Dubov, Lawrence. (2007). *Master data management and customer data integration for a global enterprise*. New York: McGraw-Hill.

Cadena-Vela, Susana, Mazón, Jose-Norberto, & Fuster-Guilló, Andrés. (2020). *Defining a Master Data Management Approach for Increasing Open Data Understandability*. Paper presented at the On the Move to Meaningful Internet Systems: OTM 2019 Workshops.

CC BY

Cleven, A., & Wortmann, F. (2010). *Uncovering Four Strategies to Approach Master Data Management*. Paper presented at the 43rd Hawaii International Conference on System Sciences System Sciences (HICSS).

Haneem, Faizura, Kama, Nazri, Taskin, Nazim, Pauleen, David, & Abu Bakar, Nur Azaliah. (2019). Determinants of master data management adoption by local government organizations: An empirical study. *International Journal of Information Management, 45*, 25-43. doi: 10.1016/j.ijinfomgt.2018.10.007

Haug, Anders, Schlichter, Jakob, Stentoft Arlbjørn, Jan, & Zachariassen, Frederik. (2013). Master data quality barriers: an empirical investigation. *Industrial Management & Data Systems, 113*(2), 234-249. doi: 10.1108/02635571311303550

Haug, Anders, & Stentoft Arlbjørn, Jan. (2011). Barriers to master data quality. *Journal of Enterprise Information Management, 24*(3), 288-303. doi: 10.1108/17410391111122862

Iqbal, R., Yuda, P., Aditya, W., Hidayanto, A. N., Handayani, P. Wuri, & Harahap, N. C. (2019). *Master Data Management Maturity Assessment: Case Study of XYZ Company*. Paper presented at the 2019 2nd International Conference on Applied Information Technology and Innovation (ICAITI).

Ishizuka, Ryo, Washizaki, Hironori, Fukazawa, Yoshiaki, Saito, Shinobu, & Ouji, Saori. (2019). *Categorizing and Visualizing Issue Tickets to Better Understand the Features Implemented in Existing Software Systems*. Paper presented at the 10th International Workshop on Empirical Software Engineering in Practice (IWESEP). Conference retrieved from

Karia, J., Sundararajan, M., & Raghavan, G. Srinivasa. (2021). *Distributed Ledger Systems to Improve Data Synchronization in Enterprise Processes*. Paper presented at the 2021 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER).

Ko, Chielsin, Adywiratama, Andytias Dwi, & Hidayanto, Achmad Nizar. (2021). *Master Data Management Maturity Model (MD3M) Assessment: A Case Study in Secretariat of Presidential Advisory Council*. Paper presented at the 9th International Conference on Information and Communication Technology (ICoICT), Information and Communication Technology (ICoICT). Conference retrieved from

Krismawati, D., Ruldeviyani, Y., & Rusli, R. (2019). *Master Data Management Maturity Model: A Case Study at Statistics Business Register in Statistics Indonesia*. Paper presented at the 2019 International Conference on Information and Communications Technology (ICOIACT).

Loshin, David. (2009). *Master data management*. Amsterdam ; Boston: Elsevier/Morgan Kaufmann.

Myers, Michael D. (1999). Investigating information systems with ethnographic research. *Communications of AIS, 2*(23), 1-20.

Radcliffe, J. (2004). Learn the Four Styles of Customer Data Integration: Gartner.

Rahman, A. Aditya, Dharma, P. Gusman, Fatchur, R. Mohamad, Freedrikson, A. Nala, Ari, B. Pranata, & Ruldeviyani, Y. (2019). *Master Data Management Maturity Assessment: A Case Study of A Pasar Rebo Public Hospital*. Paper presented at the 2019 International Conference on Advanced Computer Science and information Systems (ICACSIS).

Smith, Heather A., & McKeen, James D. (2008). Developments in Practice XXX: Master Data Management: Salvation Or Snake Oil? *Communications of the Association for Information Systems, 23*, 63-72. doi: 10.17705/1CAIS.02304

Spruit, Marco, & Pietzka, Katharina. (2015). MD3M: The master data management maturity model. *Computers in Human Behavior, 51*(Part B), 1068-1076. doi: 10.1016/j.chb.2014.09.030

Tony, E. Adams, Stacy Holman, Jones, & Carolyn, Ellis. (2015). *Autoethnography*. New York, New York: Oxford University Press.

CC BY

White, A., Newman, D., Logan, D., & Radcliffe, J. (2006). Mastering Master Data Management: Gartner.

Wilkinson, Leland, & Friendly, Michael. (2009). The History of the Cluster Heat Map. *American Statistician, 63*(2), 179-184. doi: 10.1198/tas.2009.0033

Vilminko-Heikkinen, R., & Pekkola, S. (2013). *Establishing an Organization's Master Data Management Function: A Stepwise Approach*. Paper presented at the 2013 46th Hawaii International Conference on System Sciences.

Vilminko-Heikkinen, Riikka, & Pekkola, Samuli. (2017). Master data management and its organizational implementation. *Journal of Enterprise Information Management, 30*(3), 454-475. doi: 10.1108/JEIM-07-2015-0070

Vilminko-Heikkinen, Riikka, & Pekkola, Samuli. (2019). Changes in roles, responsibilities and ownership in organizing master data management. *International Journal of Information Management, 47*, 76-87. doi: 10.1016/j.ijinfomgt.2018.12.017

Yang, Fang, Wen, Xueguo, Aziz, Asad, & Luhach, Ashish Kr. (2021). The need for local adaptation of smart infrastructure for sustainable economic management. *Environmental Impact Assessment Review, 88*. doi: 10.1016/j.eiar.2021.106565

Zúñiga, Daniel Vásquez, Cruz, Romina Kukurelo, Ibañez, Carlos Raymundo, Dominguez, Francisco, & Moguerza, Javier M. (2018). *Master Data Management Maturity Model for the Microfinance Sector in Peru*. Paper presented at the Proceedings of the 2nd International Conference on Information System and Data Mining.

CC BY