MEASURING MATURITY OF DATA-DRIVEN DECISION MAKING

Case study on selecting and applying a maturity model for a small-size professional services company

Lappeenranta–Lahti University of Technology LUT

Master's thesis in Data Analytics in Decision Making

2023

Reetu Jormakka

Examiners: Associate professor, Kalle Elfvengren

Post-doctoral researcher, Lasse Metso

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

School of Engineering Science

Industrial Engineering and Management

Reetu Jormakka

Measuring maturity of data-driven decision making – Case study on selecting and applying a maturity model for a small-size professional services company

Master's thesis

2023

66 pages, 11 figures, 3 tables, 1 appendix

Examiners: Associate professor Kalle Elfvengren, Post-doctoral researcher Lasse Metso

Keywords: maturity of data-driven decision making, data-driven decision making, maturity model, analytics maturity, data, analytics

After the rapid growth in the amount of data produced during the last ten years, organizations are heavily invested in finding ways on how to develop their decision processes into more data-driven. An effective and a widely known concept for answering this demand is the so-called maturity measurement. Maturity models provide a method for organizations to find out their current-state of data-driven decision making compared to market standards, and at the same time, a way to identify the most important development steps on their path forward.

This thesis study selects a maturity model of data-driven decision making from the existing scientific literature based on a literature review and applies it for a Finnish small-size professional services company. By using the model, a maturity level is assigned for the organization, and along the way, the most important development items for the company are identified. In addition, the thesis study provides a method on how to select the most applicable maturity model from the existing literature.

The results of the thesis study show, that selecting and applying an existing maturity model is an effective way to promote an organization's ability to move forward with their data-driven processes. The results also suggest that existing maturity models can and should be used more in the future research, instead of focusing to develop new models from scratch.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUTin insinööritieteiden tiedekunta

Reetu Jormakka

Tiedolla johtamisen kypsyystason mittaaminen – Case-tutkimus kypsyysmallin valinnasta ja soveltamisesta pienelle asiantuntijapalveluyritykselle

Diplomityö

2023

66 sivua, 11 kuvaa, 3 taulukkoa, 1 liite

Tarkastajat: Dosentti Kalle Elfvengren, Tutkijatohtori Lasse Metso

Avainsanat: tiedolla johtaminen, datavetoinen päätöksenteko, tiedolla johtamisen kypsyystaso, kypsyysmalli, analytiikan kypsyys, data, analytiikka

Viimeisen kymmenen vuoden aikana tuotetun datan määrän valtavan lisääntymisen johdosta yritykset ja organisaatiot panostavat tällä hetkellä voimakkaasti kehittääkseen päätöksentekoprosessejaan entistä datavetoisemmiksi. Eräs tehokas ja laajalti tunnettu konsepti vastata tähän tarpeeseen on niin kutsuttu kypsyysmallinnus. Kypsyysmallit tarjoavat organisaatioille menetelmän tunnistaa nykytilansa tiedolla johtamisen tasossa, ja tätä kautta auttavat niitä kehittämään datavetoisia päätöksentekoprosessejaan.

Tämä opinnäytetyö valitsee olemassa olevasta tieteellisestä kirjallisuudesta tiedolla johtamisen kypsyysmallin kirjallisuuskatsauksen perusteella ja soveltaa sitä suomalaiseen pienen asiantuntijapalveluyritykseen. Mallia käyttämällä organisaatiolle määritetään tiedolla johtamisen kypsyystaso ja samalla tunnistetaan aihepiirin tärkeimmät kehityskohteet. Lisäksi opinnäytetyö tarjoaa menetelmän oikean kypsyysmallin valintaan olemassa olevasta kirjallisuudesta. Opinnäytetyön tulokset osoittavat, että olemassa olevan kypsyysmallin valinta ja soveltaminen on tehokas tapa edistää organisaation kykyä kehittyä datavetoisissa prosesseissaan.

Tulokset viittaavat myös siihen, että olemassa olevia kypsyysmalleja voidaan ja niitä tulisi käyttää enemmän tulevaisuuden tutkimuksissa, jatkuvan uusien mallien kehittämisen sijaan.

SYMBOLS AND ABBREVIATIONS

BU	Business Unit
CEO	Chief Executive Officer
Churn	Percentage of clients that cancel or do not renew their subscription (stops using the service)
COO	Chief Operative Officer
CRM	Client Relationship Management
CRM/DM	Client Relationship Management / Delivery Management
ETL	Extract, Transform, Load (concept in data engineering, describing a process in which data is loaded from a source system, transformed or modelled into a more suitable form, and finally loaded into a destination system)
HE-BIA	Higher Education – Business Intelligence & Analytics
JSON	JavaScript Object Notation (a standardized form of structured data)
KPI	Key Performance Indicator
MRR	Monthly Recurring Revenue
REST API	Representational State Transfer Application Programming Interface (a program which is designed to enable data retrieval from information systems)
SaaS	Software as a Service
SOBIMM	Service-Oriented Business Intelligence Maturity Model

Table of contents

Appendices

Figures

Tables

1 Introduction

1.1 Background

The amount of data in the world has been exploding since the last ten years, and organizations are able to acquire and utilize more and more different data sources and types in their businesses year by year. For most of the organizations across the globe, it is not anymore a question of whether to focus on utilizing data or not, but how well is it done and how to get better at it.

To get better at something, it is usually needed to understand the current state from which to get on moving. Maturity models are a way for organizations to compare their current standing on a chosen topic against the market standards. This way organizations are able to understand the level of criticality for needed actions and possible directions for their future development programs. When it comes to the maturity of organizations' data utilization, this thesis study will use a term called "Maturity of data-driven decision making" (MDDD).

1.2 Objectives and scope

This study aims to investigate the current field of academic MDDD models and apply a maturity assessment on a selected case organization. Research around maturity of data-driven decision making is highly relevant at the moment, as even more and even smaller companies are able to become data-driven due the new tools, technologies and open data sources, on an accelerated speed. However, organizations do face challenges in their internal development, if they cannot answer the following questions: "how well are we doing with our data and analytics?" or "what are the most critical issues that our data-driven decision-making processes are experiencing?".

MDDD models are not a particularly new thing in academic literature nor in the business domain. During the last fifteen years, significant number of scientific articles have been published around the topic, resulting in as many as multiple dozens of different MDDD models developed. The models follow mostly a regular structure of assessing the maturity by scoring different organizational dimensions by maturity levels, basing the assessments on qualitative data, such as surveys and interviews.

The problem with the previous research is that it is not dozens of different kinds of MDDD models that organizations need, but instead a few well-functioning. Due to a lack of standardization and tendency to build new models based only softly on existing literature, we are seeing lot of new models built, but the field of research not really progressing. What would be needed in the field of MDDD models, would be to assess and develop the existing models with scientific methodology, by applying them for case organizations, instead of building new models every time a need for maturity measurement emerges.

The primary aim of this study is to determine the maturity of data-driven decision making for a case organization, which is a small-size Finnish professional services company. This is done by selecting an MDDD model from existing literature based on a literature review, and then applying the model for the company. By utilizing existing literature instead of building a new model for the case, the thesis study will work as a right step forward on the field of research with MDDD models, by evaluating how the chosen MDDD model can be used outside of its initial publication context. The research questions of the study are the following:

- What is the level of MDDD of the case company?
- How to select an MDDD model from the existing literature?
- How well can the chosen MDDD model be used outside of its initial case study?
- What are the main development items for the chosen MDDD model?

1.3 Execution and structure of the study

The above-described research questions are answered with a literature review on the existing MDDD research and with a case study which aims to assess whether there are existing models in the literature which can be easily used to generate value for the case study organization. The selected MDDD model is applied in the case study by leveraging a workshop session with the case study organization. The literature review works as a basis for developing the methodology to select the most applicable MDDD model. The scope of the study is tied to small size organizations and an industry of professional services. Below presented an INPUT-OUTPUT presentation of the thesis study structure (Figure 1):

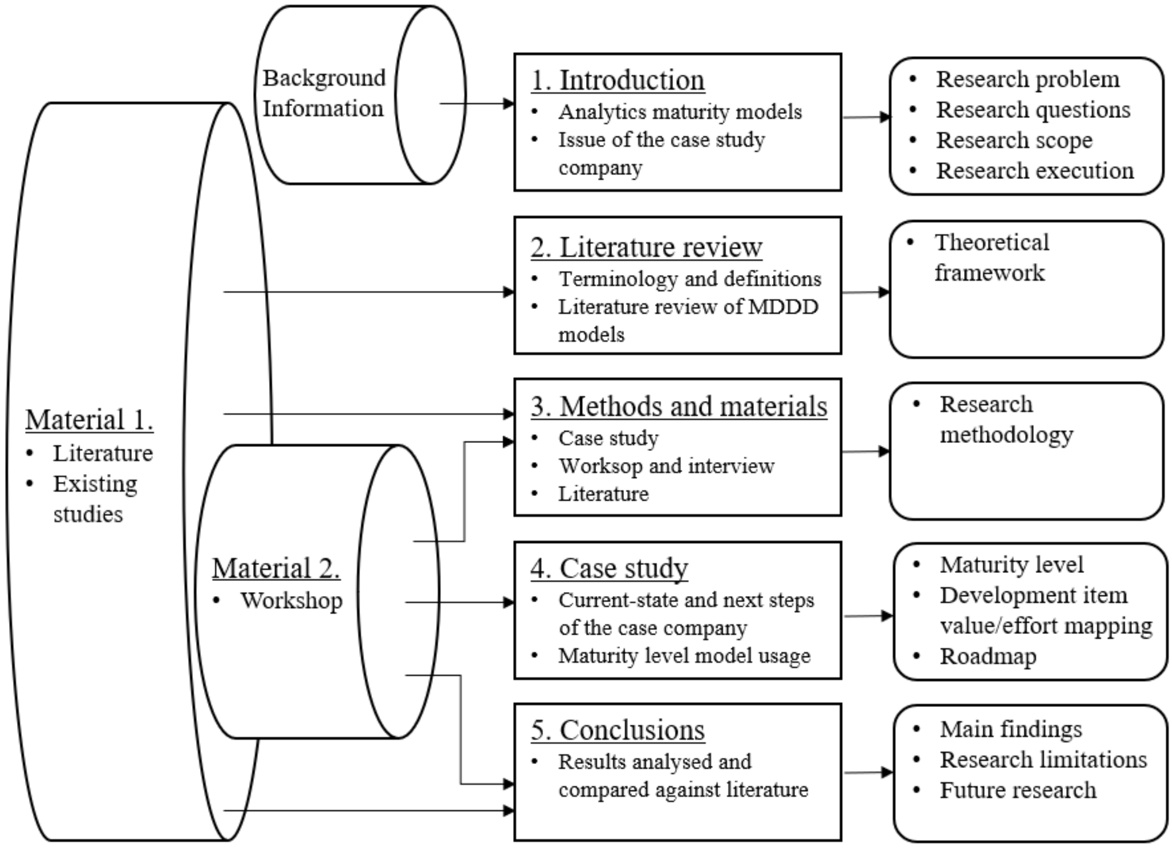

Figure 1. INPUT-OUPUT figure of the thesis study structure

Like illustrated above in Figure 1, the structure of the thesis is as follows: Theory section introduces the relevant terms and concepts, in addition to presenting a literature review on existing MDDD models. Methods section introduces the case organization and the selected MDDD model for the empirical part. Then, it walks through the process of applying the model for the case company, including the approach and interviews. Results section demonstrates the maturity of the case company, and Conclusions section includes the discussion around the results, limitations of the study and the MDDD future development.

2 Literature review on MDDD models

This chapter gives an overview regarding the theoretical framework around MDDD models. It starts by going through few essential terms and definitions around the field, goes on to a thorough literature review on different models developed already, and ends up with limitations and conclusions regarding the current state of the field of academic MDDD models, acting as a basis for the model selection for the empirical phase of the study.

2.1 Terms and definitions around maturity models for data-driven decision making

When talking about maturity models, it is important to define what is specifically being measured by them. Giving the fact that the current field of academic study around analytics related maturity models swimmingly mixes up data-driven decision making, data-driven culture, data analytics and business intelligence, it is important to define what is really meant by them, and what is therefore actually measured with the models. This chapter defines the above-mentioned terms basing on widely cited academic papers, this way preparing the reader to understand the content of the actual maturity models, which are presented later in the chapter.

2.1.1 Data-driven decision making

One of the most important concepts regarding all the data, analytics and maturity related work is the definition of data-driven decision making. Provost and Fawcett (2013) define it as follows:

> *"Data-driven decision making (DDD) refers to the practice of basing decisions on the analysis of data rather than purely on intuition."*

As expected, there is a large amount of scientific study conducted around data-driven decision making, and the benefits of basing business decisions on acquired data rather than gut feeling is widely demonstrated (Provost & Fawcett, 2013). For example, the study by Lavalle et al. (2011) shows that the tendency to apply intuition in decision making decreases the performance of an organization in all of the business process areas compared to organizations with a tendency to base decisions on data. Similarly, Brynjolfsson et al. (2011) demonstrate that organizations utilizing data in decision making are 5- 6 % more productive than organizations that do not.

The key takeaway with the definition of data-driven decision making is the decision-making process itself, and whether data is being used in it or not. It does not take into consideration e.g. how the data is analysed or which kind of technologies is being used to do so.

2.1.2 Data-driven culture

Whereas data-driven decision making as a concept focuses only on the decision processes of an organization, data-driven culture brings in wider approach, trying to tie in the cultural attitude towards data and analytics in an organizational level. For example, Berndtsson et al. (2018) define data-driven culture as follows:

> *"A data-driven culture is characterized by a decision process that emphasise testing and experimentation, where data outweighs opinions, and where failure is accepted as long as something is learnt from it."*

As we see from the definition by Berndtsson, it is not only the data-driven decision-making process, but also the related testing, experimentation and failure around it that are being accepted in an organization that has a data-driven culture. What we see in organizations with data-driven cultures, is that the data-driven culture has a positive impact on organizations' innovation and performance (Chatterjee et al., 2021). Based on the above-described definitions, one could argue that data-driven decision making is one dimension of data-driven culture of an organization.

2.1.3 Business Intelligence

When we talk about Business Intelligence, even broader and more varying concepts and definitions are available. In their book called "Business Intelligence Competence Centers – A Team to Approach Maximizing Competitive Advantage" (2006), Miller et al. define Business Intelligence as follows:

> *"Business Intelligence (BI) can be defined as getting the right information to the right people at the right time. The term encompasses all the capabilities required to turn data into intelligence that everyone in your organization can trust and use for more effective decision making."*

What we see in the definition by Miller et al., is that when comparing Business Intelligence as a term against data-driven decision making and data-driven culture, we introduce more dimensions, especially the technological one. When talking about "all the capabilities required to turn data into intelligence" we bring in systems and processes regarding e.g., data warehousing, ETL-processes, data modelling, reporting and analytics. Same goes with the definition by Elsa (2010):

> *"Business intelligence (BI) refers to computer-based techniques used in spotting, digging-out, and analysing business data, such as sales revenue by products and/or departments, or by associated costs and incomes."*

Hence, Business Intelligence as a term can be seen more technical than data-driven decision making and data-driven culture. In addition to the culture and processes around it, it aims to encapsulate the actual systems and methods on generating information from data.

2.1.4 Data Analytics

Despite its broad usage and high general awareness of it, the term "Data Analytics" is needed to be defined as well, as it plays a visible role in the data analytics maturity literature. In his book called "Data Analytics" (2016) Thomas Runkler defines data analytics as follows:

> *"Data analytics is defined as the application of computer systems to the analysis of large data sets for the support of decisions."*

What we can see from Runkler's definition is that when talking about data analytics, we are referring a quite specific operation of using computer systems to analyse data. Unlike with the other terms went through above, the term stands out quite operational, and does not incorporate organizational or specific technological dimensions (except that computer systems as a technology are needed to perform data analytics).

2.1.5 Maturity

As a final concept to be defined, we go through "maturity". The term maturity is primarily used to narrate how evolved something is, in the scope of primitive to advanced / very evolved. One of the most cited definitions of maturity is written by Simpson and Weiner to the Oxford English Dictionary (1898):

> *"State of being complete, perfect or ready"*

Another definition that represents the use of the term in the data analytics literature nicely is stated by Fraser et al. (2002):

> *"To reach a desired state of maturity, an evolutionary transformation path from an initial to a target stage needs to be progressed."*

Applying the definition by Fraser et al., (2002) to field of data analytics and data-driven culture, an organization that has just started on the path of data analytics and therefore lacks the experience and fails to carry out data analytics initiatives, can be stated as an organization with a low data analytics maturity, whereas an experienced organization that is successfully leveraging data and analytics in its everyday business with a state of the art methodology, can be stated as an organization with a high data analytics maturity.

Regarding the measurement of maturity of something, there are few important characteristic of scientific maturity models: maturity concept, dimensions, levels, maturity principles and assessment (Lahrmann & Marx, 2010). The maturity concept states the actual objective which's maturity is measured (like data-driven culture), whereas dimensions and levels state different aspects that need to be measured (like BI-systems, decision processes, data assets) and the levels of assessment (like low, medium, high). The maturity principle and the assessment state that what is the method of assigning different levels for different dimensions (like continuous or staged, and qualitive or quantitative).

2.1.6 Conclusions on terms and definitions around analytics maturity models

By going through the most important terms and definitions within the MDDD literature, this chapter aimed to prepare the reader to acknowledge that the scientific articles around MDDD are filled up with terms and concepts which tend easily to be left undefined in the original papers. However, with this short literature review done in this chapter, we can lock the meanings for the rest of this study. Based on the definitions and articles stated above in the chapter, the definitions for the terms above in the context of this thesis study are as follows:

- Data-driven decision making: A practice of using data in the decision processes wherever possible, rather than intuition.
- Data-driven culture: An organizational culture that focuses to leverage data in its decision making by accepting testing, experimentation and failure around data all over the organization.
- Business Intelligence: A set of systems and processes to enable an organization to gather, store, analyse and utilize data in its decision-making processes.
- Data Analytics: An operational and systematic way of analysing data to derive information from it.
- Maturity: State of being ready, in the scope of primitive to highly evolved

2.2 Reasons to measure maturity of data-driven decision making

Before taking the actions and investing into measuring a maturity of data-driven decision making with an MDDD model, one might consider why to do this, and not just focus on the organization’s analytics in their own way without regarding and comparing to others. What does it benefit to know how mature you are?

According to Bititci et al. (2015), there are two aspects how measuring maturity benefits organizations in their performance. Firstly, using a maturity model framework encourages the management team to talk openly regarding the matters that have been measured with the model. This helps everyone in the organization to get more involved and take ownership of the assessment's results. It also makes it easier for the organization to learn and improve. This, in turn, helps the management team get better at what they do and become more critical of their practices, which strengthens the organization's ability to learn and grow.

Secondly, using a maturity model approach speeds up the process of getting assessment results. This makes it quicker and easier to review the organizational ways of working. As a result, it encourages more frequent check-ins to see how things are going. This helps the organization keep learning and getting better, and it also supports the ongoing development of managerial skills. In other words, development is difficult to make happen, if there is no knowledge of the fact that something needs to be developed.

Similarly than with Bititci, Cech et al. (2018) states that with the usage of a data competence maturity model, an organization is better capable of executing data-driven business decision making, by being better able to asses data usage and develop data capabilities. Only by understanding the current state of things (which is what maturity models provides), is an organization able to take actions on the right development items, using the right volume of investment. The key takeaway for organizations of using MDDD models is the identification of the required steps to move from one maturity level to the next. This makes it easier to select and prioritize the actions that are the most potential to deliver impact.

One way to approach the way how a maturity measurement generates value is to consider a maturity measurement as a gap analysis. Gap analysis is a systematic method used to assess the variance or disparity between a current state (the existing situation or performance) and a desired future state (the intended or optimal condition) within an organization or a process (Kim & Ji, 2018). Gap analysis involves evaluating the differences in processes, practices, capabilities, or performance metrics. By identifying these gaps, whether they relate to skills, resources, processes, or objectives, organizations gain insights into what needs to be improved, developed, or changed to bridge the divide between the current and desired states. Gap analysis serves as a roadmap, guiding decision-making and enabling organizations to prioritize actions and allocate resources effectively to close these gaps, ultimately facilitating progress toward desired goals and objectives. With an MDDD model, the gap analysis is performed by placing the case organization onto a maturity level, and then identifying the needed improvements to reach the next level. For instance, Gökalp et al. (2022) demonstrates a data science maturity model, which combines the concepts of a maturity model and a gap analysis.

There are lot of evidence showing success gained by using maturity models. By referring to Cech et al. (2018) the utilization of maturity models in both small and large organizations across diverse sectors has notably enhanced their operational processes. Adoption of these models for process development has resulted in significant enhancements across several key areas such as cost efficiency, scheduling accuracy, task completion times, overall productivity, quality standards, customer satisfaction, and return on investment. According to Goldenson and Gibson (2003), the application of data analytics competency maturity models in has yielded remarkable successes. For instance, Boeing achieved a 50 percent reduction in release times, while General Motors substantially increased the rate of meeting development milestones from 50 to 95 percent. Additionally, Lockheed Martin realized a notable 30 percent increase in software productivity. These success stories illustrate significant improvements attained by these organizations by utilizing maturity models (Gibson et al., 2006).

2.3 Maturity models for data-driven decision making

During the last 20 years, dozens of studies have been conducted to develop different kinds of models for measuring the maturity of data-driven decision making, data-driven culture, data analytics and Business Intelligence. This chapter contains a literature review of the existing models, focusing on five selected papers. As there are lot of models developed in the field throughout the years, the selection of these five particular papers is done by aiming to represent the variety in the field by choosing different kind of models for the review. While focusing on the five selected, many other papers are referred in addition.

2.3.1 LaValle's transformation oriented 6 dimension – 3 stage maturity model

One of the most cited analytics maturity models in academic setting is LaValle's (LaValle et al., 2011) 3-stage model with 6 dimensions, published in 2011. The study by Lavalle et al. bases on a large 3000 respondent survey conducted by MIT and IBM to gather information about data usage of different organizations across the globe. The results of the survey show significant correlation with organization's analytics usage and its performance, and that decisions in especially some management areas are more easily left for intuition than based on data. For example, finance and strategy are based on at least some data even among the low performers, whereas general management and workforce management tend to get left on intuition.

The study classified all the organizations from the survey into three different categories, based on their analytics maturity: aspirational, experienced and transformed. These were then set as the stages of the maturity model. Aspirational level stands for the lowest maturity.

In the aspirational level, organizations focus on their existing and traditional data to cut costs of their existing processes. In the experienced level, organizations start to look for new ways to gather and utilize data in their organization, once they have recognized the positive effects of the aspirational phase. In the transformed level, analytics and data are key differentiators of organizations and they utilize data and analytics all over the organization (LaValle et al., 2011).

The six dimensions (row headers) of the model were based on the survey questions, classifying the different answers and states that the organizations stated in the survey. Hence, the model is basically built by putting together all the survey responses into a matrix (Figure 2), and then ordering the answers into three stages based on the level of complexity of each feature.

Three capability levels — Aspirational, Experienced and Transformed — were based on how respondents rated their organization's analytic prowess.

	ASPIRATIONAL	EXPERIENCED	TRANSFORMED
Motive	•Use analytics to justify actions	•Use analytics to guide actions	•Use analytics to prescribe actions
Functional proficiency	•Financial management and budgeting •Operations and production •Sales and marketing	•All Aspirational functions •Strategy/business development •Customer service •Product research/development	•All Aspirational and Experienced functions •Risk management •Customer experience •Work force planning/allocation •General management •Brand and market management
Business challenges	•Competitive differentiation through innovation •Cost efficiency (primary) •Revenue growth (secondary)	•Competitive differentiation through innovation •Revenue growth (primary) •Cost efficiency (secondary)	•Competitive differentiation through innovation •Revenue growth (primary) •Profitability acquiring/retaining customers (targeted focus)
Key obstacles	•Lack of understanding how to leverage analytics for business value •Executive sponsorship •Culture does not encourage sharing information	•Lack of understanding how to leverage analytics for business value •Skills within line of business •Ownership of data is unclear or governance is ineffective	•Lack of understanding how to leverage analytics for business value •Management bandwidth due to competing priorities •Accessibility of the data
Data management	•Limited ability to capture, aggregate, analyze or share information and insights	•Moderate ability to capture, aggregate and analyze data •Limited ability to share information and insights	•Strong ability to capture, aggregate and analyze data •Effective at sharing information and insights
Analytics in action	•Rarely use rigorous approaches to make decisions •Limited use of insights to guide future strategies or day-to-day operations	•Some use of rigorous approaches to make decisions •Growing use of insights to guide future strategies, but still limited use of insights to guide day-to-day operations	•Most use rigorous approaches to make decisions •Almost all use insights to guide future strategies, and most use insights to guide day-to-day operations

Figure 2. Analytics maturity model by LaValle et al., 2011

As seen in Figure 2 above LaValle's model presents quite a lot of verbal descriptions of the different levels, which makes it easy and fast to interpret, present and use.

2.3.2 Service-Oriented Business Intelligence Maturity Model

Based on a solid literature review on different business intelligence maturity models, such as HP's model (HP, 2007), TDWI:s model (Eckerson, 2009) and IBM's model (IBM, 2009), Shaaban et al. (2011) have developed Service-Oriented Business Intelligence Maturity Model (SOBIMM) which aims to overcome pitfalls of the previous models, such as ability to tie the model and the metrics into the organization, and its processes and services (Figure 3).

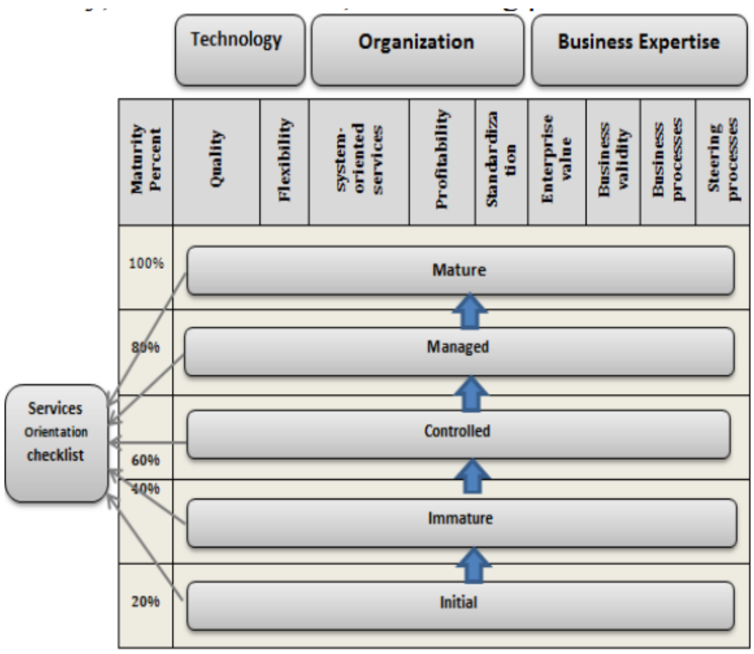

Figure 3. SOBIMM model by Shaaban et al. (2011)

SOBIMM model works in a way, that the set of questions that are used to define the maturity, is called "Services orientation checklist". The services orientation checklist is gone through with all the relevant business stakeholders, in order to gain understanding from each of the model metrics (quality, flexibility, system-oriented services, and so on). The metrics are classified into three different categories, technology, organization and business expertise. Each metric is given the maturity between "Initial" and "Mature", which then gives the overall maturity of the organization regarding business intelligence (Shaaban et al., 2011).

2.3.3 Buitelaar's extensive 10-dimension organizational maturity model

One of the most mathematical and complex analytics maturity models is the model developed for Mobiel.nl by Ruben Buitelaar (Buitelaar, 2018). Buitelaar's model is based on large literature review by compounding dimensions and stages from several existing models and developing some scoring logic on top of them. Buitelaar builds his model based on the following existing models:

- TDWI Analytics Intelligence Maturity Model (Halper & Stodder, 2014)
- Maturity Model for Business Intelligence and Performance Management (Hostmann & Hagerty, 2010)
- The HP Business Intelligence Maturity Model (HP, 2007)
- Big Data Business Model Maturity Chart (Schmarzo, 2016)
- Big Data & Analytics Maturity Model (Nott, 2014)
- Adobe Analytics Maturity Model (Rigby & Contreras, 2015)
- The Five Stages of Analytical Maturity (Davenport & Harris, 2017)

The model by Buitelaar has 10 dimensions through which the analytics maturity is assessed: Data, metrics, skills, technology, leadership, culture, strategy, agility, integration and empowerment. The assessed organization is walked through a list of questions, each question allocated for one dimension, and scored between 0 and 5. As there are different number of questions for different dimensions, the resulting points per dimension are weighted on the number of questions of each dimension. This results as a numeric score from 0 to 10 for each dimension.

By using the numeric scores, the model classifies the assessed organization into one of the five different maturity stages, the stages being: Reporting, analyzing, optimizing, empowering and innovating. Similarly with the model by Lavalle (Lavalle et al., 2011) the stage descriptions show how low maturity organizations are focused on reporting and optimizing existing legacy, while organizations with high analytics maturity are innovating and creating new with analytics. Below we can see the Buitelaar's model applied (Figure 4), demonstrating average scores and scarcity of the answers based on from 14 respondent organizations.

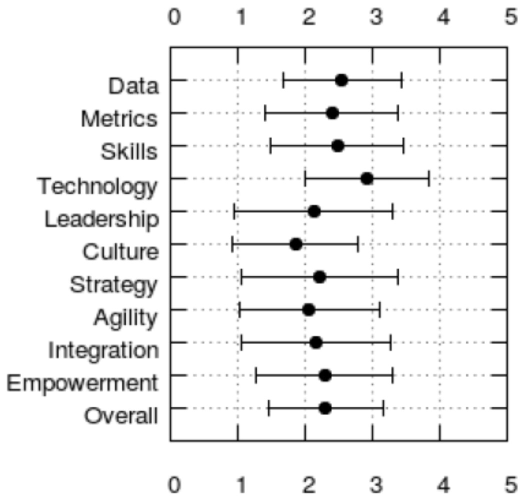

Figure 4. Maturity model by Buitelaar (2018)

What we see from Figure 4 above, is that scores averaged from 14 organizations are mostly between 2 and 3, which is expected. However, "Technology" raises higher than the others, whereas "Culture" is the lowest dimension in average. These initial results by Buitelaar (2018) raise an interesting question, whether companies tend to have the needed technology in place for mature data-driven decision making, but the organizational culture doesn't yet enable its full-scale utilization.

2.3.4 Industry specific "Higher Education Business Intelligence and Analytics" maturity model for higher education

In addition to a vast amount of general-purpose analytics maturity models which are designed to be used in a wide range of organizations and domains, there are also several industry specific models, such as the HE-BIA (Higher Education – Business Intelligence and Analytics) maturity model for higher education by Elsa and Xiaomeng (2022) and the analytics maturity model for health care by Brooks et al. (2013). With an industry specific model, an organization is able to take more detailed industry specific metrics into consideration when defining the analytical maturity, than what might be possible with a general-purpose model.

Cardoso et al., (2013) explain, that in the field of higher education, the analytics maturity assessment of an organization is particularly difficult because the lack of understanding the main analytics use cases of the field, in addition to finding the right professionals in organizations that are capable of answering the wide range of maturity assessment questions. As an answer for this issue, Elsa and Xiaomeng (2022) developed a so called HE-BIA maturity assessment model (Higher Education Business Intelligence and Analytics Maturity Model), which aims to make the maturity assessment easier regarding the above-mentioned challenges.

Like many others, also the HE-BIA model is based on a wide literature review on existing analytics maturity models, from which the main dimensions and levels are derived. However, to make the model more relevant for the field of education, many of the business-related terms and concepts are converted into educational context. As much as 18 different dimensions were chosen for the model, which makes the HE-BIA one of the most dimensional maturity models available. The 18 dimensions are classified into 7 categories, which are value, program/project management, business process/ BI&A development, people, data, data products and technical foundations. When it comes to levels, HE-BIA has traditional five level maturity ranking. Below (Figure 5) we can see all the dimensions through a sample organization from the study by Elsa and Xiaomeng (2022).

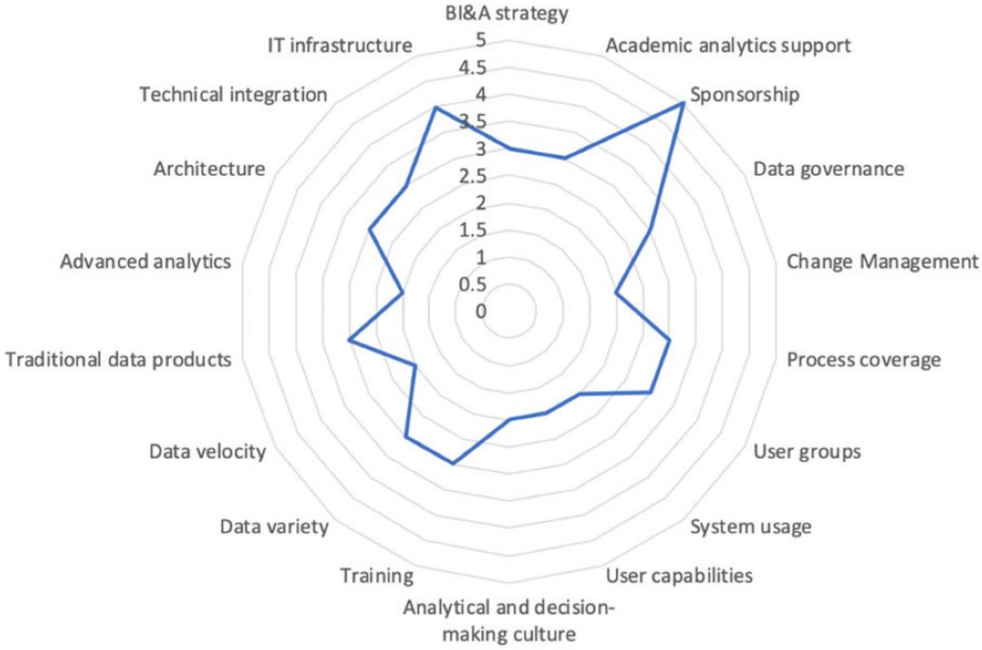

Figure 5. HE-BIA maturity model by Elsa & Xiaomeng (2022)

As we see in Figure 5 above, the sample organization analyzed in Elsa's and Xiaomeng's paper represents an organization with a quite high analytics maturity regarding sponsorship and IT infrastructure, but quite low maturity regarding advanced analytics and change management, for instance. The spider chart visualization type chosen by Elsa and Xiaomeng represents nicely the strengths and weaknesses of the sample organization, giving directions for future actions.

2.3.5 Maturity model of being a data-driven organization

One of the most cited analytics maturity researchers from the last few years is Mikael Berndtsson from the university of Skövde. He was part of the research team of an article called "Analyzing Business Intelligence Maturity" (Gudfinnsson et al., 2015) and then few years later leading a study described in an article called "Becoming a data-driven organization" (Berndtsson et al., 2018).

In the study by Berndtsson et al. (2018), a framework of enabling factors of a data-driven organization is conducted by interviewing a case organization's employees and external consultants and presented in a form of a mind map. The aim of the map is to represent all the dimensions and relationships between factors, regarding all the important things that an organization needs to consider when planning to become more data driven (Figure 6).

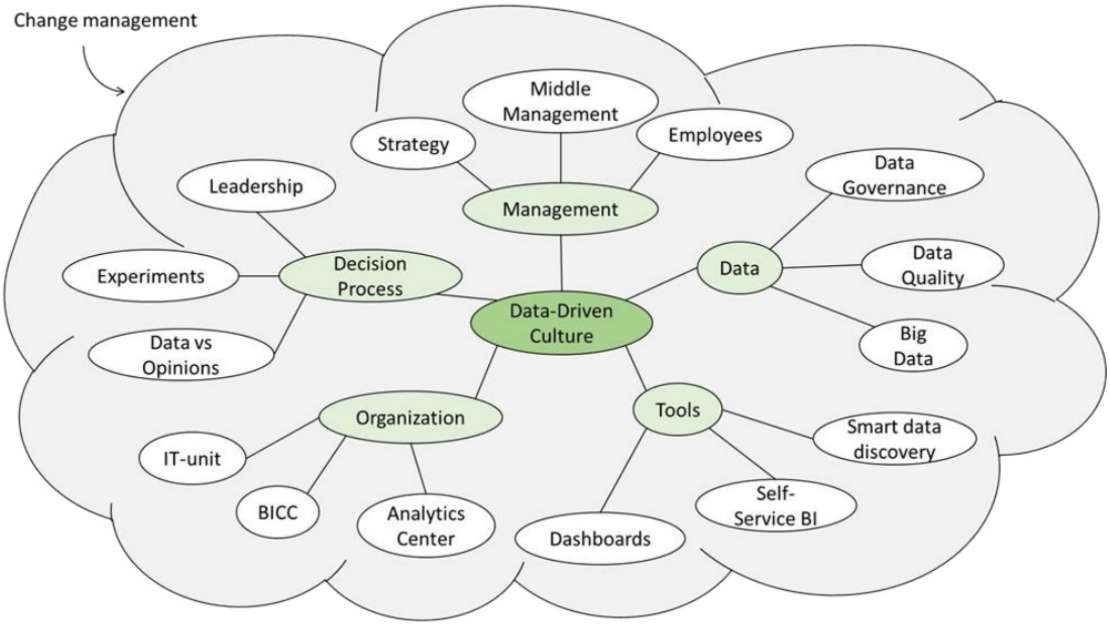

Figure 6. Mind map of enabling factors for implementing a data-driven culture Berndtsson et al. (2018)

From the mind map presented above (Figure 6), we can see how the enabling factors of a data-driven organization can be classified into five different categories, the categories being Management, Data, Tools, Organization and Decision Process. Even though we can see similarities regarding these categories with several other maturity models (Buitelaar, 2018; Shaaban et al., 2011), the categories are especially similar with the analytics maturity model by Elsa & Xiaomeng (2022). Similarly, there are a lot of overlap regarding the actual dimensions between these models.

In addition to a framework of enabling factors of becoming a data-driven organization, Berndtsson et al. (2018) represent a first version of an actual maturity model for assessing the level of it, called Analytics capability maturity model (Figure 7).

	Level 1	Level 2	Level 3	Level 4
Organisation	No explicit BI or analytics unit	A dedicated BI unit is established	BI and advanced analytics are separate units	An organisation wide analytics team is established
Technology	Mostly spread-sheets	Data warehouse is in place	DW and data mining tools are being used	Insights are opera-tionalized ASAP
Decision Process	Hippo-culture	Reports and dash-boards are generated automatically and on demand	Test and learn culture	(semi) Automatized decisions
People	Little trust in data and ana-lytics	Mixed feelings about analytics	Self-Service DW Mixed feelings about advanced analytics	Self-Service Analytics
Analytics	Descriptive	Descriptive	Descriptive, predictive	Descriptive, predic-tive, prescriptive

Figure 7. Analytics capability maturity model by Berndtsson et al. (2018) used for a case organization

What is different with the model by Berndtsson et al. (2018) to many other maturity models, is that it bases less on direct guidelines on how to calculate and define the maturity score on each dimension. Instead, it provides a relatively simple matrix of four levels, into which organization is to be mapped after a qualitative consideration of organization's enabling factors (Figure 6). However, the five dimensions of the matrix are not directly coming from the framework of the enabling factors, but instead from their separate analytics capability model.

2.4 Conclusions on the literature review of MDDD models

As a conclusion on the literature review on MDDD models, we see that they serve as comprehensive roadmaps for organizations, enabling them to systematically evaluate their existing data capabilities and progress toward more advanced and effective analytics practices (e.g. Król and Zdonek, 2020; Bititci et al., 2015; Cech et al., 2015). The models offer a structured assessment of an organization's strengths and weaknesses in handling data, allowing for a detailed understanding of where improvements are needed. By providing a clear trajectory of maturity levels and associated benchmarks, MDDD models guide strategic planning, resource allocation, and capability development efforts. As companies advance through these maturity stages, they gain the ability to make more informed decisions, mitigate risks, streamline operations, and ultimately drive innovation, giving them a competitive advantage in leveraging data for enhanced performance and sustainable growth in a data-driven business environment.

What we see from the maturity models went through in this chapter, is that the main maturity model characteristics (Lahrmann & Marx, 2010), like stated already in the chapter 2.1.5, are very clearly visible. We see that all the models above are built on dimensions through which the assessment is made, and levels, which indicate a certain level of maturity for each dimension. Then based on the scores of each dimension, the overall maturity level of an organization is assigned.

In addition to similarities within the structures of the models, we see clear homogeneity between the dimensions of the models. Like stated already in the chapter 2.2.5, the dimensions of pretty much all the models cover one way or another at least data, tools, management/leadership, technology and people/culture related aspects. Hence, we can conclude that by taking account at least these few high-level dimensions for a MDDD model, a comprehensive view on an organization’s data-driven decision making can be accomplished.

A literature review by Król and Zdonek (2020) provides a comparison table, which demonstrates how 11 different MDDD models align on their high-level maturity level definitions (Table 1). Their study demonstrates the overlap between the existing literature in the field of MDDD models, and how different models can be used for similar results. In Table 1 below, we see 11 different MDDD models selected and compared by Król and Zdonek (numbers from 1 to 11, model references in Appendix 1), each presented on their own rows. Then, in the columns numbered from 1 to 6, we see different maturity levels that are available in each of the models. For example, the lowest level maturity, level 1, is described as “Building reports” in the APMM model (Grossman, 2018) and as “Standalone

Analytics" by Logi Analytics (2019). Even though differences can be seen, many of the level descriptions on the same levels are very similar with each other (Król and Zdonek, 2020).

Table 1. Analytics maturity model stage comparison table by Król and Zdonek (2020)

AMM	A Stage in the Analytics Continuum				
	1	2	3	4	5
1	Building reports	Building and deploying models	Building and deploying analytics	Enterprise-wide processes for analytics	Analytics is strategy driven
2	–	–	–	–	–
3	Laggard	Follower	Competitor	Leader	Innovator
4	Learning	Planning	Building	Applying	Leading
5	Analytically Impaired (Not Data Driven)	Localized Analytics (Use Reporting)	Analytical Aspirations (See the Value of Analytics)	Analytical Companies (Good at Analytics)	Analytical Competitors (Analytical Nirvana)
6	Basic	Opportunistic	Systematic	Differentiating	Transformational
7	Standalone Analytics	Bolt-On Analytics	Inline Analytics	Analytics Infused	Genius Analytics
8	–	–	–	–	–
9	Analytically Unaware	Analytically Aware	Analytically Astute	Empowered	Explorative
10	Nascent	Pre-Adoption	Early Adoption	Corporate Adoption	Mature Visionary
11	Impaired Initiated	Operational	Integrated	Competitor	Addicted

1. Analytic Processes Maturity Model (APMM). 2. Analytics Maturity Quotient Framework. 3. Blast Analytics Maturity Assessment Framework. 4. DAMM—Data Analytics Maturity Model for Associations. 5. DELTA Plus Model. 6. Gartner's Maturity Model for Data and Analytics. 7. Logi Analytics Maturity Model. 8. Online Analytics Maturity Model. 9. SAS Analytics Maturity Scorecard. 10. TDWI Analytics Maturity Model. 11. Web Analytics Maturity Model. Source: own elaboration.

References of the comparison table at Appendix 1.

The main differences that we see currently between the existing MDDD models are related to the scope and objective of the models (e.g. Elsa's higher education oriented domain specific model (Elsa & Xiaomeng, 2022) versus LaValle's general purpose model for all domains (LaValle et al., 2011), in addition to the varying backgrounds of the entities behind the models (e.g. private companies creating models to boost their sales (Gartner, IBM, HP…) versus scientific studies conducted by researches and universities). When it comes to differences between the models, it should also be noted, that basically every single model comes with its own questionnaire and way of colleting the data of an assessed organization.

2.5 Limitations with MDDD models

Despite the significant number of different MDDD models developed over the last fifteen years, there are few persistent issues that seem to hold on to the field, even though new models emerge year after year.

The first problem is the number of the models themselves. While having a high number of different models, some with academic background, some with backgrounds from consultancy companies, and something from between, it is easy to end up in a situation where there are too many models to create an environment with comparable measurement results. Like stated for example by Lasrado et al. (2015), the field of MDDD models lack standardization. For the future of scientific MDDD models and their usage, it would be beneficial to guide organizations and researchers to use and developed same models, in order to standardize their usage and the models themselves, instead of developing a new model every time a new need to assess maturity emerges.

Another widely recognized issue within existing MDDD model literature is the lack of theoretical background behind the model development (e.g. Ariyarathna & Peter, 2019; Lasrado et al., 2015; Cosic et al., 2012). What we see in the field of MDDD is that lot of models, especially on the non-academic side, tend to come up for example with the model dimensions without basing them on legitim research results. Hence, the right direction for the future work regarding MDDD models would be to focus on utilizing the existing work done and basing the model development on peer reviewed scientific publications.

Related to the lack of theoretical background, lot of MDDD models seem to miss valid documentation (Brooks et al., 2015) as well. This can be seen also by going through the publications of the models represented in the literature review of this study (chapter 2.2). While some of the papers do state the methodology behind the model development and instructions on how to use them (e.g. Buitelaar, 2018), most of the papers do not explicitly open up for example the questionnaire behind the data acquisition (e.g. Shaaban et al., 2011). What we see is a negative self-reinforcing cycle when models are published without legitim theoretical background or proper documentation, and therefore new models are needed to built from scratch despite large number of existing work.

As the data analytics field rapidly evolves with emerging technologies and trends, existing maturity models often face challenges in keeping pace with these advancements (Dutta et al., 2022). The inability of old models to incorporate and address the characteristics of these new developments can lead to limitations in accurately capturing the diverse dimensions of modern data analytics practices. Consequently, this limitation might hinder organizations from fully leveraging the potential benefits offered by cutting-edge technologies or adapting swiftly to transformative changes within the data landscape.

2.6 MDDD model comparison table

Like stated in the chapters 1, 2.4 and 2.5, there are plenty of maturity models in the context of academic literature. On top of that, the bulk of them even happen to be relatively similar with each other. That is the reason why this thesis study demonstrates a comparison table for MDDD models, aiming to provide help for organizations and case studies in the future to select a proper model for their use case.

Majority of the existing academic literature uses MDDD model comparison tables as a basis for creating new models or purely for feature comparison (e.g. Ariyarathna and Peter, 2019), instead of providing guidance for selecting an existing one for a case. Some existing comparison tables do consider the actual usage of the models into some extent (e.g. Król & Zdonek, 2020), but not a major focus on the selection or usability is considered. With the comparison table presented in this chapter below, this study hopes to steer the future development of MDDD models into the direction of continuing the existing work, instead of always developing a new model for each need, stated as one of the development items in the field by Lasrado et al., (2015).

The comparison table presented in this chapter will not incorporate all the possible options from the existing literature, but instead showcase the usage of the table with the MDDD models presented in the chapter 2.3. For this particular thesis study, the whole literature was explored throughout a literature review, from which these five potential ones were chosen. The comparison table will showcase, how the MDDD model selection for the thesis case study is done between these five models, which were already qualified from the existing literature.

When compared to MDDD comparison tables demonstrated by Król & Zdonek, (2020) and Ariyarathna & Peter, (2019) the comparison table developed by this thesis work case study focuses more heavily into practical usability of the models, as that has been identified as one of the main issues in utilizing the existing work of the field (e.g. Ariyarathna, & Peter, 2019; Lasrado et al., 2015; Cosic et al., 2012). The chosen dimensions for the MDDD model comparison table are listed below, along with the reasoning for their selection:

- Questionnaire documented: Majority of the MDDD model academic papers do not showcase how to survey or interview the case study organizations in order to point out their maturity scores
- Other usage documentation: Many papers do not point out any guidance on how to apply their model for other cases
- Applicability for the domain: As shown in the chapter 2.3, some MDDD models are only applicable for specific industries
- Relevancy of dimensions: Even as Król & Zdonek, (2020) are counting the number of dimensions in different MDDD models, the relevancy of the dimensions for the case organization is more relevant from the model usability point of view
- Measurement methodology: Some models are simpler and some more complex than others. This can affect the usability of the model, depending on the case and the case organization

The resulting table, built based on the above-mentioned arguments, is presented below (Table 2).

Table 2. MDDD comparison table for the thesis case study, filled in with theoretical point scales of each column

Model name	Questionnaire documented	Other usage documentation	Applicability for the domain	Relevancy of dimensions	Measurement methodology	Totals
LaValle et al., (2011)	0 - 5	0 - 5	0 - 5	0 - 5	0 - 5	**0 - 25**
Shabaan et al., (2011)	0 - 5	0 - 5	0 - 5	0 - 5	0 - 5	**0 - 25**
Buitelaar (2018)	0 - 5	0 - 5	0 - 5	0 - 5	0 - 5	**0 - 25**
Elsa & Xiaomeng (2022)	0 - 5	0 - 5	0 - 5	0 - 5	0 - 5	**0 - 25**
Berndtsson et al., (2018)	0 - 5	0 - 5	0 - 5	0 - 5	0 - 5	**0 - 25**
Totals	**0 - 25**	**0 - 25**	**0 - 25**	**0 - 25**	**0 - 25**	

Scoring: from 0 to 5 for each cell (0 as lowest possible and 5 as highest possible)

For demonstration purposes, the comparison table is here filled in with theoretical point scales of each column. The selected models are the ones presented in the chapter 2.3. Each of the actual scoring cells for models per column range from 0 (the lowest possible score) to 5 (the highest possible score). The measures' main objective is to demonstrate the relative differences between the models, not define absolute and global scores for the models that could be applied outside of the selected model group. The Totals-column demonstrates the highest scoring model, whereas the Totals-row demonstrates the dimensions that are performing better than others when considering all the models.

3 Methods and materials

This chapter presents the methodology and materials of the thesis study. As the thesis is based on a case, the chapter starts by assessing case study as a research method, and then moves forward into the methods used in the case, them being an interview and a workshop session. In addition, the materials used in the workshop session are presented.

3.1 Case study as a research method

A case study is a research method that forms a methodological approach characterized by a comprehensive analysis of a specific instance or phenomenon within its real-life context (Yin, 2009). For example, Grimaldi, et al. (2019) study the correlation between advanced data utilization and business performance by applying a regression model for a real-life organization, which works as a research environment of the study.

Case study methodology involves an intensive examination of a single entity or a bounded system, aiming for comprehensive understanding rather than statistical generalization (Heale & Twycross, 2018). By emphasising real world use cases, the method enables researchers to bring in observations and opportunities for validation, that could not necessarily have been possible to acquire in a controlled research environment. In other words, a case study shows the distinction between what happened and what was found, in contrast to what was initially planned (Anderson & Arsenault, 2005). The main strength of the case study method is the strong connection between theory and practice, in addition to its holistic nature and ability to combine multiple sources of evidence (Gummesson, 2000). It is well suited for studies which require input from the real world (Yin, 2009), such as when a theoretical model is applied against a real-world organization (Grimaldi et al., 2019). The main limitations of the case study method are related to the generalizability of the case studies, as often the case study results are not easily applicable in other circumstances (Johnson, 1995).

In this thesis study, the case study methodology is used to apply an existing MDDD model for a real-life organization (the case study organization) in order to verify, how well an existing maturity model can be used outside of its original publication context, how well the maturity of the case study organization can be measured with the theoretical model and what are the main development items and future research needs for the model and for the field in general, based on the findings. The chosen method is well aligned with the case study method literature, when it comes to the most suited methods for the research questions of the thesis study (Grimaldi, 2019; Yin, 2009; Gummesson, 2000; Johnson, 1995)

3.2 Workshop and interview as research methods

A workshop is a structured meeting for carefully selected and relevant persons, to follow the meeting agenda and discuss, collaborate, work together, refine, define and reach closure on a selected topic (Ørngreen & Levinsen, 2017). The main objective of a workshop is straight forward: to produce data about the question in hand. The participation group is typically kept small, which allows all the participants to actually participate in the workshop activities, which is also expected of them. Based on the level of the workshop preparation, workshops can be classified into three categories:

- Low level preparation: The workshop follows predefined guidelines, such as topic and format of discussion, but no major predefined activities are placed (e.g. Chambers, 2002).
- Mid-level preparation: The workshop is prepared with phases, activities, roles and progression (e.g. Müllert & Jungk 1987).
- High-level preparation: The workshop with phases, activities, roles and progression is ran through with additional dimensions, such as scenarios and critique rounds (e.g. Soneryd & Amelung, 2016).

The main benefits that are achieved with workshop methodology are tied to the active participation of people, which is enforced by the nature of a workshop. With active participation, low hierarchy and low mental barriers to speak up, workshops enforce creativity and innovation, in addition to drawing out information, that might have been left out in a more formal set up (Öberg & Hernwall, 2016). However, workshops do have limitations regarding the documentation and reliability of the acquired data (Ørngreen & Levinsen, 2017).

Interview is a qualitative research method, in which a researcher asks questions from a research participant in order to acquire data. The three widely known interview types are structured interviews, semi-structured interviews, and unstructured interviews (Wilson, 2012). In a structured interview, the same set of questions is asked from all of the research participants, and no room for additions or discussion is given. A typical example of a structured interview is a survey. In a semi-structured interview, the researcher has some room to move, for example as a form of guiding questions. In an unstructured interview, the flow of the interview is often like a discussion, in which the researcher guides the discussion into a certain direction to acquire the needed information, with no predefined questions placed. As a research method, an interview is most suitable in research contexts, where human issues or experiences are studied, or when the research questions are best answered with verbally instead of numbers (Beck & Manuel, 2008).

In this case study, workshop methodology is applied to acquire information regarding the case study organization’s current data assets, information technology selections and data-driven processes and issues. The case study workshop follows the guidelines of a low-level preparation workshop (Chambers, 2002), by incorporating elements of an unstructured interview (Wilson, 2012). Even though the research questions are not entirely the most suitable domain of qualitative methodology (Beck & Manuel, 2008), the method selection is based on the fact that no documentation exists regarding the data, technology and processes of the case study organization, and no easy access can be given to the company systems to gather data. Therefore, an interview-like workshop is the most suitable method to gather a holistic current-state view of a wide scope of things, in a short time period. In addition, workshop methodology enables activities such as value/effort mapping with the case study organization representative. The practical step-by-step run through of the workshop is presented in the chapter 4.3, (process flow chart presented in Figure 8).

3.3 Materials

As presented in Figure 1, the main material of the thesis study is the existing literature, formed into a literature review (chapter 2). The literature review provides a view into the existing academic MDDD models, this way enabling the selection and application of the most suitable model for the case study.

This thesis case study did not utilize data sets or documentation in the case study, that would have been delivered or gathered outside the workshop phase, as the case study organization didn’t have applicable material to be used or handed over for research purposes. However, a written documentation was produced as a part of the workshop, describing data assets, information systems, organization, decision processes and management regarding data-driven decision making of the case study organization. In addition, few case study organization’s information systems were explored as part of the workshop, to see and validate parts of the discussion in practice. The content of the generated documentation is described in detail in chapter 4.4. Then, the material that was created as a part of the workshop, was then utilized in the maturity level assignment, value/effort mapping and roadmap creation in the latter parts of the workshop (illustrated in Figure 1 and described in detail in chapter 4).

4 Case study

In this chapter, the case study and its results are presented. The chapter starts by representing the case study organization. After that we go through the selection process of the MDDD model which was used in the maturity definition of the organization. Then, the actual results of the maturity measurement along with the methods used are presented. In addition to the actual maturity measurements, the thesis study provides a roadmap based on the maturity measurement, for the case organization to start on the path of developing their maturity of data-driven decision making.

4.1 Case study organization

The case study organization of this thesis study is a Finnish small-size marketing company. Regarding its age, the company can be considered as a start-up. It employs few dozen employees.

The business model of the case study organization is to provide professional services in the field of marketing for other Finnish companies. The services are delivered and invoiced in an hourly rate manner. For the sold hours, the marketing specialists of the case study organization take care of particular marketing activities of the client organizations. Hence, the company does not possess other major assets than its specialists, whose working time is sold for the client organizations in exchange for the responsibility of taking care of the agreed tasks for them. In addition to the actual operative client work carried out by the marketing specialists, the company runs sales, client management, HR and marketing functions supporting the client work.

The case study company is mainly managed by CEO (Chief Executive Officer) and COO (Chief Operative Officer) roles, the first one being responsible of sales and finance, while latter manages the chargeable client work. The mid-level management layer is relatively thin, making the COO both strategically and operatively key stakeholder in the sense of data utilization in running the company. In the context of this case study, the COO of the case study organization is identified to play the best possible role in understanding the data utilization in the company: both in the C-level strategic decision-making, as well as in the day-to-day staffing and employee related matters.

Background reasoning for the case study company to take part into a study of maturity in data-driven decision making is related to the doubts of the company management regarding proper utilization of data and analytics in the organizational decision making. Regarding the young age of the company, the top management has not had time to take on initiatives to start developing systems that would enable data gathering and analytics. In addition, considering the nature of professional services business (gathering data from industrial machines might often be more straight forward than gathering data from people), it has been felt difficult by the top management to take action on starting the development. However, the company is now eager to find out their current MDDD and pinpoint the most valuable development items based on the assessment.

4.2 MDDD model selection

Based on the characteristics of the case study organization presented in the previous chapter, this chapter demonstrates how the comparison table developed by the thesis study (Table 2) is used to select an MDDD model for the case study organization. The case study organization characteristics that affect the resulting scoring the most, are its size (revenue, operations, personnel, etc.) and its industry (professional services). These characteristics highly affect the “Relevancy of dimensions” and “Measurement methodology” scores, whereas the “Questionnaire documented”, “Other usage documentation” and “Measurement methodology” scores are mostly given based on the model characteristics, regardless of the case study organization characteristics. Below, the comparison table is filled and presented, in the context of the case study (Table 3), after which the arguments behind the scores are given.

Table 3. MDDD comparison table for the thesis case study, filled in with the scores for MDDD models of the chapter 2.3

Model name	Questionnaire documented	Other usage documentation	Applicability for the domain	Relevancy of dimensions	Measurement methodology	Totals
LaValle et al., (2011)	0	1	3	3	2	**9**
Shabaan et al., (2011)	0	0	1	2	1	**4**
Buitelaar (2018)	2	3	3	4	5	**17**
Elsa & Xiaomeng (2022)	0	2	0	3	3	**8**
Berndtsson et al., (2018)	3	2	5	5	3	**18**
Totals	**5**	**8**	**12**	**17**	**14**	

Scoring: from 0 to 5 for each cell (0 as lowest possible and 5 as highest possible)

As we can see from Table 3 above, the best questionnaire documentation between the five models can be found from the model by Berndtsson et al. (2018), as the framework which defines the elements of the organization that are needed to go through for defining the overall maturity level is presented (Figure 6). However, Berndtsson's score reaches only 3/5, as the framework leaves a lot of room for interpretation (which however might be beneficial in some cases as well). In contrast, Buitelaar (2018) provides some actual questionnaire questions as well but utilizing them flexibly with the thesis case study organization is difficult. Other models do not provide any material besides the final maturity matrix. As we can see from the bottom-level Totals-row, the "Questionnaire documented" was the most poorly performing dimension of these five models, acquiring only 5 points in total.

Regarding other usage documentation, most of the models are documented poorly, which is also stated as a known issue in the field by Brooks et al. (2015). Buitelaar's work is best documented, which is natural given the length of the publication. It is several times longer than the other papers.

Assessing the applicability for the domain is quite straight forward, at least with the model by Elsa & Xiaomeng, which is designed for another industry than the one in which the thesis case study organization operates. On top of the actual industry, some models are clearly developed for larger and for already more mature organizations than the thesis case study organization. For these reasons, Berndtsson's model is the most suitable, giving the most flexible and easy-to-adapt framework for small size companies as well.

Relevancy of dimensions was the best performing dimension in the comparison table by scoring 17 points in total, regarding these five assessed models. All the models gave relatively numerous and relevant dimensions. The dimensions by Shaaban et al., (2011) were a bit difficult to understand, whereas the dimensions by LaValle et al., (2011) should have been a bit more detailed for an easy use. Then again, dimensions by Elsa & Xiaomeng (2022) were a bit too detailed, as all the aspects presented were not relevant regarding the thesis case study organization.

When it comes to measurement methodology, most of the assessed MDDD models did not really have any sophisticated scoring process, but instead the maturity level was assessed based on the maturity matrix level description of each cell. For example, the description of level 3 maturity of "organization" dimension in Berndtsson's (2018) model is described as "BI and advanced analytics are separate units", and then it is the assessors' job to find out whether this is the case or not regarding the organization whose maturity is under assessment. Even though the method is relatively simple, it does not necessarily make it bad, as a model that it easy-to-use and understand, and not too specific, is also easy to apply for different kind of use cases. However, the highest score of the category goes for the model by Buitelaar (2018), as his model is quantified with a continuous variable (not maturity classes, but a continuous maturity score, which enables more accurate results and a slightly more scientific approach), in addition to his thorough statistical validation of the methodology.

From the Totals-column of the comparison table we see that the two clearly most suitable MDDD models for assessment of the thesis case study organization are the models by Berndtsson et al. (2018) and Buitelaar (2018). Surely both could have been easily used to generate proper results, but as the model by Berndtsson et al. (2018) resulted with a slightly higher total score, that is the model that was selected for the thesis case study.

4.3 Case study execution

At November 2023, the empirical case study workshop was carried out with the representative (COO) of the case study organization. This chapter demonstrates how the workshop was carried out. Below, a flow chart of the case study process is demonstrated (Figure 8), after which each step is described in more detail.

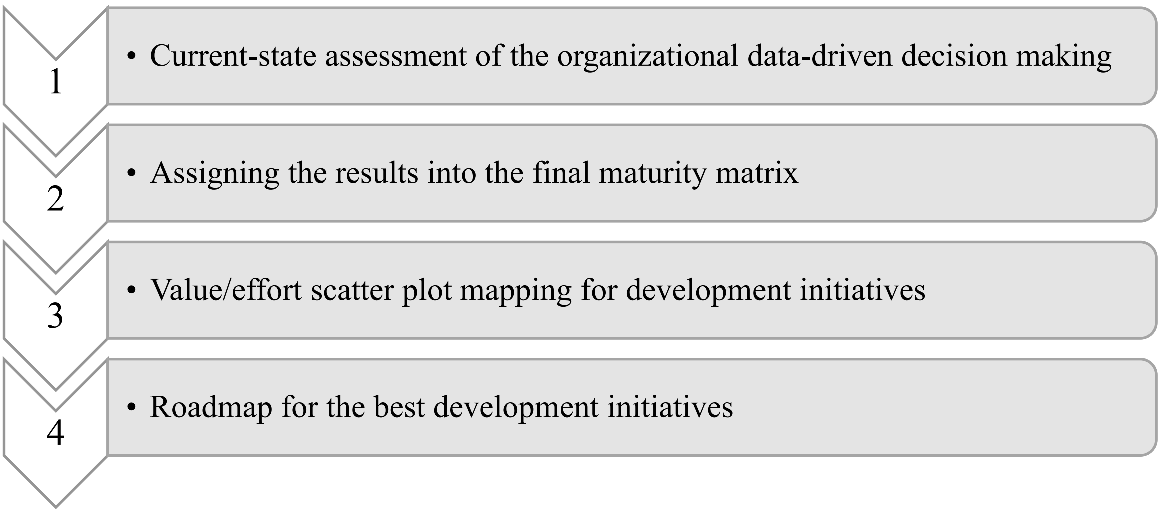

Figure 8. Case study execution process chart

Case study execution process chart steps described:

1. Current-state assessment of the organizational data-driven decision making: First, Berndtsson's (2018) framework of enabling factors (Figure 6) for data-driven culture was used as a basis for discussion. The aim of the step was to use the framework to walk through all the relevant aspects of an organization, regarding data-driven culture and decision making. This step was carried out in an interview-like manner, going through the enabling factors step by step, and asking the COO of the case study organization to give a current-state overview of the enabling factors from the viewpoint of the case study organization. The discussion was not strictly limited to the framework's enabling factors, but instead the framework was used as a guide for the discussion, ensuring that all the relevant aspects were taken into consideration. In addition to the organizational current-state of the enabling factors, data-driven decision processes of the organization and currently experienced issues regarding these matters were discussed as well. In addition to the discussion, the step 1 incorporated exploration of the CMRM/DM tool and the current data warehousing solutions in practice. As a result of this step, a written memo of current data assets, information systems used, data gathering, storing and utilization processes, currently

experienced issues and development items, as well as other discussion regarding the enabling factors, was generated out of the raw notes of the discussion.

2. Assigning the results into the final maturity matrix: After the discussion based on the enabling factors of the Berndtsson's (2018) framework, a relatively accurate view on the case study organization's data analytics infrastructure and operations could be drawn, as described in the previous chapter (step 1). As the dimensions in the framework of the enabling factors (Figure 6) were in line with the dimensions of the actual maturity matrix (Figure 7), the discussion based on the framework of the enabling factors provided the needed information for placing the case study organization into the matrix quite easily. Hence, as a result of this step, the case study organization was located on to the Berndtsson's maturity matrix, which then provided the overall maturity level information of the case study organization's data-driven decision making.

3. Value/effort scatter plot mapping for development initiatives: As described in the chapter 2.2 (Bititci et al., (2015), Cech et al., (2018)), it is the things around knowing your organizations maturity level that generate competitive advantage, not just the actual maturity level. For example, by going through the concepts around organization's data-driven decision making (step 1) and defining the maturity level for the organization (step 2), it becomes much easier for the organization to identify and prioritize development items for the future. Hence, in the step 3, the discussion points from the step 1 were formulated into development items with the COO of the case study organization, which were than mapped on to a scatter plot with axes of value and effort (Figure 10, Results-section). By mapping the development items in such way, it is easier for the organization to review which development items bring in the most value with relatively least effort. As a result of the step 3, a scatter plot with the best five development items enabling better data-driven decision making were visualized based on their effort and the value that they should bring.

4. Roadmap for the best development initiatives: After identifying and visualizing the most important development items, the last part of the workshop with the case study organization was to break down the development items into smaller tasks and place the development items and tasks on to a roadmap, which visualizes the timeframe for their implementation. By using a roadmap for mapping the future tasks on to a

timeline, it is easier for organizations to demonstrate the upcoming work for the organization's personnel and for possible stakeholders. It makes it also easier to allocate resources for implementation, in addition to demonstrating possible causalities and dependencies between development items and tasks.

4.4 Results

This chapter demonstrates the results of the case study. The structure of the chapter will follow the flow of the case study execution process chart (Figure 8), starting out by demonstrating the current state of the data-driven processes of the case study organization, and then moving on to the maturity level assignment. After that, the identified development items regarding the data-driven processes of the case study organization are mapped on to a value/effort matrix, and finally all the way into an implementation roadmap.

4.4.1 Current state of data driven decision making per dimension

Data

The data assets of the case study organization are mainly tied into its personnel and clients, as professional services businesses often do not involve for example machinery, supply chains, subcontractors (other than maybe other professional services companies), physical products, R&D, spare parts, product lines, physical facilities, equipment or other "data generators", as for example many industrial organizations do.

The operative data with which the case study organization is mostly dealing with consists of hourly rates, hours billed, hours used, work time usage, utilization and feedback. With these data assets, the company manages and steers the main operative business, which is the hourly work sold and done for the clients. By analysing and reporting the amount of sold work per professional, and comparing it realized timesheets, in addition to compensation information per professional, it is possible to review and guide the efficiency, gross margin, resource allocation and resource availability of the operative client work, which is crucial for the business. The most important actual data points in the operative domain are the allocated and realized hours per client per professional/role (who is doing, for who, how much, and what was done). Along with simple structured tables demonstrating the work amounts sold per client per professional, the actual realized hours ("the timer data") is gathered by using a timer to record the number of minutes spent with each client's activities. The feedback data plays also a significant role in the operations of the case study organization, as feedback given by the client is usually one of the main performance indicators of a professional regarding their client work. In addition to feedback given by clients to the case study organization, the performance of the professionals is also evaluated internally in a peer review format, which then complements the feedback given by the client, forming together a holistic view on the professional's performance.

The other highly important data utilization area for the case study organization is the sales. Active sales pipeline data gathering around marketing activities provides data around potential clients, potential client contact points, client contact information, opportunities, efficiency information of the marketing activities, in addition to actual contractual data points after closed deals, which are for example hours to be delivered and the hourly rate (which are then utilized also in the operations management). Information like client's current placement in the pipeline, opportunity size, size of the client's revenue, business area of the client, earlier contacts and opportunity source are also used to enable more efficient B2B-sales activities when planning for the next steps with a client or an opportunity. In addition to actual sales, many of the above-mentioned data points are also utilized in the client management after a closed won, when the client as an opportunity is retitled as a client in the process of continuous services, and when the client starts to generate monthly reoccurring revenue (MRR).

In addition to the operative data and sales data, the case organization gathers and utilizes regulatory and relatively general administrational data, like most companies. This includes personnel data, like general personal information, payment information, contracts, compensation model and payments, holidays, hours worked (for managing overtime), and sick leave days. Similarly with the finance, regulatory information regarding accounting and audit, like billing, revenue, gross margin, payments and cost structure is gathered and managed. In addition, recruiting and applicant information is gathered.

The most significant issue identified regarding data quality is the fact that not every professional remembers to always use "the timer" when performing client work. This makes it significantly more difficult to make pricing decisions with the timer data, as it cannot fully be trusted, without additional actions to find out the quality of the data.

Tools

The are a handful of critical information systems used by the case study organization, along with a couple of dashboards and few spreadsheets for analytics. From the operations point of view, the most relevant tool combines client relationship management (CRM) and delivery management functions, and which therefore generates the bulk of the data that is used in the operative management of the case study organization.

The CRM/delivery management tool (CRM/DM) comes to play when an opportunity is marked as a closed won in the sales pipeline management tool. At this stage, the case study organization prepares to start the monthly delivery of the service for the client, and client information from the sales tool is manually copied to the CRM/DM tool. In addition to basic client information, such as client name and contact details, also contract details are inserted, such as hours and services sold. In the CRM/DM the sold service is allocated for a certain professional, who then starts to deliver the service on a monthly basis. The professional's main use case for the CRM/DM tool is "the timer" which is put on whenever the professional starts to execute activities for the client, and put off, when the professional is finished. This way the tool provides data not only regarding the sold services for the client, the MRR based on the contract, and who is the allocated professional to deliver the service, but also regarding what was the realized amount of work done for the client, which might differ from the amount of work sold. In addition to the timer, the CRM/DM tool provides task management interface for the professionals, to organize and prioritize their activities with different clients.

Before client data is inserted into the CRM/DM tool and a deal is closed, a sales pipeline tool is used by the sales staff and managers in the organization to manage data and activities around sales. The main difference in managing client specific data with the sales pipeline tool and the CRM/DM tool is the client life cycle. When a deal is closed, the client is "transferred" into the CRM/DM tool from the sales pipeline tool. Prior to that, the sales pipeline tool is the master for managing basic client information and all the sales and marketing activities targeted to them. Here, opportunity sizes, client meeting dates and notes, client stages in the pipeline and sources of opportunities are managed. In addition, the sales pipeline tool is integrated with an open data source providing client specific revenue and marketing information, which can be utilized when designing an approach to contact the opportunity.

In addition to the CRM/DM tool and the sales pipeline tool, there are handful of other administrative information systems used in the case study organization. The accounting service of the case study organization comes with an accounting system, which provides the key financial figures for the company in monthly basis, such as revenue and gross margin. HR-tool stores all the personnel related information, from which the compensation data is the most important from operative point of view. In addition to compensation information, the HR-tool stores personnel contact details, contracts, working hours, overtime and sick leave days. For a professional, the HR-system comes with an interface for submitting the weekly timesheets (worked hours) for overtime and payment calculation.

Case study organization's data warehousing and analytics are carried out with cloud-based spreadsheets and a cloud-based data visualization/reporting tool by a global cloud platform company, in addition to an outsourced reporting service, which comes with a browser-based reporting interface. There are two main spreadsheets used for analytical purposes, into which data is manually transferred from the above-mentioned information systems, mainly from the CRM/DM tool. The spreadsheets contain dozens of sheets, which act as database tables, and the reporting tool is directly connected to these sheets. Any possible ETL-transformations (Extract, Transform, Load) that the data might need during the process of flowing from the source systems into the spreadsheets and forward to the reports, are manual. No analytical calculations or modelling is performed during the process, except simple measures (variables divided by other, etc.) in the reporting end.

When it comes to data warehousing with spreadsheets, there is an issue of data integrity. Spreadsheets are vulnerable of being accidentally used wrong, in addition to being easily exposed to typos, accidental data loss or truncations, or data type mixes. In addition, the content of the two main data warehousing spreadsheets do overlap, even though their data pipelines are not connected or synchronized, which means, that there is a risk of managing the same information manually in two different locations, and there for exposing their content to differentiate.

Organization

The way how the case study company is organized around analytics is relatively simple. From the internal staff, it is mostly only the COO who works with internal data and analytics in the every-day work. The COO is the system owner of the CRM/DM tool, which is the most data intensive tool in the organization and the most significant data generator. The COO owns the data of the CRM/DM tool and is responsible of its quality and gathering processes. In addition to the actual CRM/DM tool, the COO also manually keeps another of the main data warehousing spreadsheets ("operations data set") up to date, by ingesting data from the CRM/DM tool, the HR-tool and the sales pipeline tool into it. The operations data set brings together clients, sold work amount and services, MRR, professionals and teaming, which enables decisions around resourcing and performance.

The only actual data analyst role of the case study organization is outsourced for a company which focuses on providing a data gathering and reporting service around their SaaS-service (Software as a Service). The process around the reporting SaaS-service works in a way, that a named professional from the SaaS-service providing company is allocated to manually gather and bring in case study organization's data into a specific spreadsheet ("reporting data set"). The allocated analyst is provided with an access to the case study organization's information systems, from which the analyst manually transfers data into the spreadsheet once a month. Then with some configuration, the SaaS-software visualizes the data gathered, and the report is provided for the case study organization as a service with a monthly fee, including the cost of the analyst gathering the data. Unlike the dashboard built internally by the case study organization (based on the "operations data set"), this report by the SaaS-service is mostly based on the financial data of the case study organization, hence, its data being gathered from the accounting system. However, also data from the CRM/DM tool is brought in and combined with the financial data. Main figures that are followed with the SaaS-report, are monthly net sales revenue, gross margin and working capital, in addition to the figures built based on the CRM/DM tool, such as costs and incomes per teams and clients.

The most critical issue with the SaaS-reporting system is the reporting of financial figures based on the accounting data. The problem with the accounting data is naturally that it is only after a month when the monthly accounting is performed. When considering the delays that come in when the accountant creates the calculations and provides the data, and when the outsourced analyst takes the accounted financial data in to the data warehouse and from there to the SaaS-report, it might often be that the report represents the situation of one to two months back in time. Hence, the case study organization is not able the follow the current state of things from the SaaS-report. In addition, the SaaS-report does currently not support drilling into enough details, which would be necessary for resourcing, performance and root-cause analyses.

Decision Process and Management

The relatively small size and young age of the case study organization, combined with the professional services industry in which it is operating, have created quite difficult starting points for the company to be data-driven in its internal decision making. However, the amount of data and the number of data-driven processes are increasing year by year, which is seen as a positive trend in the company. Referring to the COO of the company, the organization and its personnel have absolutely nothing against data, but utilization of it is not yet where it’s supposed to be, as the company has just started on the path of developing their internal processes, after focusing on making the actual business stable during the first years of the company. Even though there are still decision made on intuition, especially around sales, there are already several data-driven decision processes in the organization, both on strategic and operational level, which are described here below.

Starting from the strategical level, the company strategy is yearly analysed based on multiple data sources. Client analytics, conducted mostly by the CEO of the case study organization, analyse the current client portfolio regarding their MRR, “client convenience” (as in whether the client is easy-going and easy to serve, versus difficult to work with), titles of the contact persons, services offered and churn (the measure of when and how many clients stop using the service). By aggregating and cross referencing these factors against each other, there are often insights found on what type of clients are most valuable, and there for where to aim in the future sales.

Another analytics case in the strategy work is the market size modelling and competitor analyses. Only by comparing to others is an organization able to really tell whether they are performing or not. The market size is modelled by selecting a pool of competitors from the shared market and following their annual revenue. If the revenue by the case study organization is lagging from the market trend (average of the competitors’ annual growth in revenue), the company is internally labelled as performing behind the market. Similarly, if the annual growth in revenue by the case study organization exceeds the average of competitors, the company is labelled as growing faster than the market. This approach gives the company a clear reference point on the yearly performance and works as an objective in the annual target setting.

On the operative level, there are several data sources and KPI: s (Key Performance Indicator) that are used in the everyday management of the client work and professionals. The most important and widely used KPI is utilization. For the case study organization, utilization is a measure of efficiency, calculated by dividing the hours worked with the clients (the number of billed hours) and the theoretical total employee working hours (158 monthly, 37,5 weekly). The resulting measure describes, what is the percentage of work time that goes into serving clients (direct value add), versus the percentage of working time that goes for supporting functions or waiting for work. Theoretically, in the regard of the professionals working with the clients (excluding sales, marketing and internal management) the higher the utilization, the more productive the business.

Another KPI used in the client work management of the case study organization is called the match rate. The match rate is calculated by dividing the number of sold hours with the actual hours that it took the perform the client activities. The resulting measure describes how well a professional was able to perform the tasks that were sold and planned, and at the same time, how realistic was the pricing and the work effort estimation performed in the sales stage. In an ideal situation, the sold work is performed using the estimated hours, or faster. Theoretically, the faster the activities, the more productive the business.

Utilization and the match rate are widely used in resourcing, sales planning and performance management of the case study organization. New clients are handed over for teams and professionals, whose current utilization is low, which means that they have the best availability to take in more clients. However, the COO of the case study organization identified, that they should possibly complement this approach by applying internal pricing regarding professional seniorities (an experienced and talented professional is able to deliver more value with 100 % utilization than not experienced). In addition to resourcing, utilization is used in performance management and as a basis of the company’s compensation scheme. Utilization can be assessed on several organizational levels, including a single professional, a team, a business unit (BU), or the whole operative staff of the company. The higher the utilization of a professional, the higher the performance, and therefore the higher the possibility for a higher compensation. Same applies for the team and the BU level.

In addition to utilization, also the match rate is used in the everyday performance management of professionals. An example of this is a recently established data-driven process, where the match rate is reviewed on monthly basis in a supervisor check in, and if the rate falls out of a threshold of 80 % - 120 % (hours used against hours sold), a closer look is given on what is causing the inefficiency of this particular professional. For cases where the match rate falls below 80 % (less than 80 % of the hours sold are actually used to perform the work for the client), a closer look is given for checking for example that could it be that the client is underserved.

Despite the usage of utilization and the match rate as KPI: s in the organizational management, their monitoring and reporting is not yet where it should be. The current spreadsheet-based data warehouses and the third-party reporting tool do not enable sufficient real time monitoring and drilling down from BU: s to single professionals regarding the measures. This makes it manual, inaccurate and not systematic to make and management decisions based on the data.

An example of the case study organization’s process where data and analytics would be needed, but none is yet to be used or implemented, is recruiting. There is a constant need to balance the amount of work force in relation to new won client engagements, but also at the same time keep the cost levels under control. In situations, where a sales peak has not been overseen in advance, there might be difficulties to provide professionals to take over the sold work. From growth perspective, this is a very undesired situation, as the clients that cannot be served, are forced to seek the needed services from competitors. However, from the cost structure point of view, it is not possible to keep professionals in “reserve”, waiting for surprising sales peaks, as that is not financially profitable nor motivating for the professionals. In an optimal situation, internal operations data and external market demand data could be combined, in order to come up with a forecasting model to alert from an upcoming resource demand. With help of a demand forecasting model, work force could be arranged in a more optimal manner.

4.4.2 Maturity level of the case study organization

After going through the organization, technology, data, decision processes and management of the case study organization, following the guidelines of the framework of enabling factors of data-driven culture by Berndtsson et al., (2018) (Figure 6), it is possible to place the case study organization on to the maturity matrix (presented in the chapter 2.3, Figure 7) and assign the MDDD level based on the workshop discussions with the company representative. The maturity of data-driven decision making of the case study organization is presented below (Figure 9):

	Level 1	Level 2	Level 3	Level 4
Organization	No explicit BI or analytics unit	A dedicated BI unit is established	BI and advanced analytics are separate units	An organization wide analytics team is established
Technology	Mostly spreadsheets	Data warehouse is in place	DW and data mining tools are being used	Insights are operationalized ASAP
Decision Process	Hippo-culture	Reports and dashboards are generated automatically and on demand	Test and learn culture	(semi) Automatized decisions
People	Little trust in data and analytics	Mixed feelings about analytics	Self-Service DW Mixed feelings about advanced analytics	Self-Service Analytics
Analytics	Descriptive	Descriptive	Descriptive, predictive	Descriptive, predictive, prescriptive

Figure 9. The maturity level of data-driven decision making of the case study organization, presented through the maturity matrix by Berndtsson et al. (2018).

As presented above, we can see that the case study organization locates mostly on the level 2 stage in terms of maturity of data-driven decision making. Regarding the "Organization" dimension, the case study organization exceeds the level 1 by having an explicit resource for BI and analytics, the resource being the third-party SaaS-analytics/reporting service. However, the level 3 remains unreachable, as no advanced analytics is in place at the company. Given the size of the company, it would not be reasonable to require the BI unit to be internal or consist of several professionals, and therefore one external resource can be seen as a dedicated BI unit.

In terms of technology, the case study organization does indeed use mostly spreadsheets (which is described as Level 1 characteristic in the matrix) but uses them in a manner that can be seen as data warehousing (data is explicitly gathered from several source systems to one combined location for purposes of reporting and analytics (Ponniah, 2004)). However, the level 3 stage of technology “DW and data mining tools are being used” remains clearly out of reach, as no technology designed explicitly for the purpose is utilized.

Regarding the dimension of “Decision Process”, the case study organization falls down on to the level 1. Even though the company does utilize dashboards and reporting in its every-day management, the issues identified regarding the drilling and real time monitoring capabilities of the reports are too significant for making it possible to say that the company would be able to create insights and dashboards on demand. In addition, visualized or analysed data is used mostly by the top management of the company. However, considering the utilization of several KPI: s, few above mentioned data-driven decision processes (chapter 3.4.1) and some dashboards, it is safe to say that the level 2 is almost there.

When it comes to “People” dimension, the case study organization is placed to level 2 once again, mainly because the level 3 is out of reach due the lack of advanced analytics and analytics focused tools, such as a relational data warehouse technology. From the people perspective this can be seen as a cultural issue as well, as major steps have not yet been taken to implement workflows and technologies for the purpose. However, as described in the chapter 3.4.1, the personnel of the company are down for analytics, but the overall volume and focus just are not there quite yet, as the company has been established still quite recently, and the analytical workflows haven’t taken a significant role yet. And, yet several data analytics related steps forward have been already taken during the recent couple of years.

The case study organization’s maturity level regarding the “Analytics” dimension is placed on to the level 2. This is quite straight forward, as no predictive analytics has been implemented yet. However, as mentioned in the chapter 3.4.1, the case study organization has identified the need for predictive market demand estimation.

As a conclusion, we can see the case study organization being placed on to the level 2 maturity in all except one dimension (Figure 9). This means that in overall the case study organization is assigned to the level 2 maturity. In general, we see lot of initiatives, potential and future use cases for more advanced analytics use cases, but not yet that much implemented, which clearly lifts the company above the level 1, but still leaves a visible gap to the level 3.

4.4.3 Development item value/effort mapping

Like presented in the chapter 2.2. (Bititci et al., 2015; Cech et al., 2018) the real benefits for an organization from using an MDDD model realize as an updated current-state understanding and as identified development items regarding data analytics. Hence, after assigning the actual organizational MDDD level for the case study organization, the case study continued by assessing the most valuable development items for the company, based on the comprehensive current-state understanding gained through the workshop discussion with the company COO.

The chosen methodology for finding out first stage development items in the terms of data and analytics for the case study organization was to find out which identified issues from the step 1 discussion would bring in the most value, and at the same time, be feasible to implement with least effort. This was done by mapping the most significant development items on to a scatter plot (Figure 10) with a vertical axis of effort and a horizontal axis of added value. The development items were placed on to the scatter plot with the case study company COO "intuition and discussion based", meaning that no quantitative assessment was used to place the items to their locations. Hence, the main idea in Figure 10 is to demonstrate the relations between different development items, instead of actual levels of effort or added value. A scatter plot designed this way forms four different zones (a quad chart), which were classified as follows:

- **Zone A**: High effort & low added value: These items are waste of resources and should not be moved forward, as they demand a lot of work to make them happen, but do not generate much added value.
- **Zone B**: High effort & high added value: These items are heavy to make happen, but as a reward, they do generate lot of added value. Should be moved forward, but as their resource consumption is high, careful planning is needed.

- **Zone C**: Low effort & low added value: These items are easy to implement, but no real impact is going to take place due to the low potential of added value. Might be usable to move forward, but not with priority.
- **Zone D**: Low effort & high added value: These items, sometimes also referred to as "the low hanging fruits", are easy to implement but do still generate lot of added value. Should be treated as priority one development items.

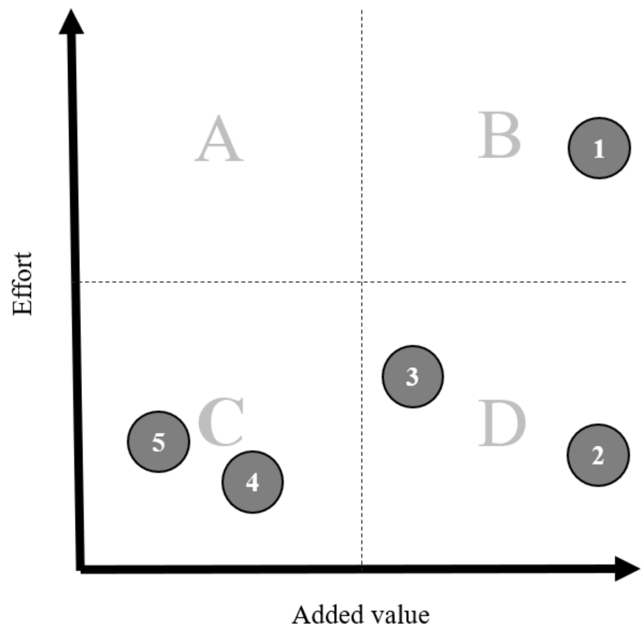

Figure 10. Value/effort mapping of the case study organization's development items

The above presented development items are:

1. Deployment of a relational database (cloud) as an analytics data warehouse and building of automated data pipelines from the CRM/DM, the sales system and the accounting system into it, with scheduled ETL-processes.
2. Deployment and configuration of a dashboard for operational management, which is directly connected to the data warehouse (development item number 1), and which enables drilling down from BU level KPI: s all the way into individual professional level.

3. Planning and implementation of an incentive process for operative professionals to remember to generate “the timer data” regarding client work activities.
4. More active managerial level intervention into intuition-based decision making, especially around sales.
5. Development of a market demand estimation model, combined with data regarding own capacity, for implementation of a data-driven process of alerting for future recruitment need.

As we can see from Figure 10 above, the most relevant development initiative for the case study organization would be to deploy a relational data warehouse for analytics, along with the few needed data pipelines and some reporting tied into it (development items 1 & 2). Building the pipelines would be the more labor-intensive part (located in the zone B), and then putting up the report referring to the structured, modelled and up-to-date data would be the easier part (zone D). The ETL-pipelines for the data to be transferred from the source systems into the data warehouse would require the source systems to provide their data through REST API: s (which they do), and then a cloud-hosted and scheduled script to fetch the data from the API: s and ingest them into the data warehouse, along with some possible transformations/modelling included. However, some reporting tools are also already able to be connected directly into JSON REST API: s (e.g. Microsoft, 2023), which would mean that no separate fetching script would be needed.

Along with these two above-mentioned items, we can see the development of an incentive process for the timer data generation in the zone D as well, which indicates it to be an easy and effective target. Development of an incentive process for timer data generations means, that the case study organization needs to come up with a process, which “forces” the professionals to use the timer when doing client work. This would result in a more whole and accurate set of data, which would then lead into more accurate decisions. An example of a process like this would be for instance to integrate the timer data input as a mandatory part of the weekly working hour reporting.

In the zone C we can see the two least effective, but on the other hand low effort development items; more active managerial level intervention into intuition-based decision making, and development of a market demand estimation model. Intervention into intuition-based decision-making would probably naturally take place due to other initiatives around data and analytics, whereas the predictive recruitment model would require some modelling work. Considering the current data sources for the model, it might be difficult to deliver real value that would actually make a difference in the current recruitment process. However, when the maturity of the internal data and its usage increases, also the predictive modelling becomes more relevant.

4.4.4 Roadmap for implementation

As there are certain dependencies between the development items, and as resources are always limited, it is useful to plan the implementation of different initiatives in advance. For this purpose, a simple Gant chart was drawn as a part of the case study workshop with the company COO (Figure 11). The purpose of the Gant chart is to demonstrate, when do certain implementation activities take place, what should be the order of the activities, and is it possible to proceed with multiple development items at the same time.

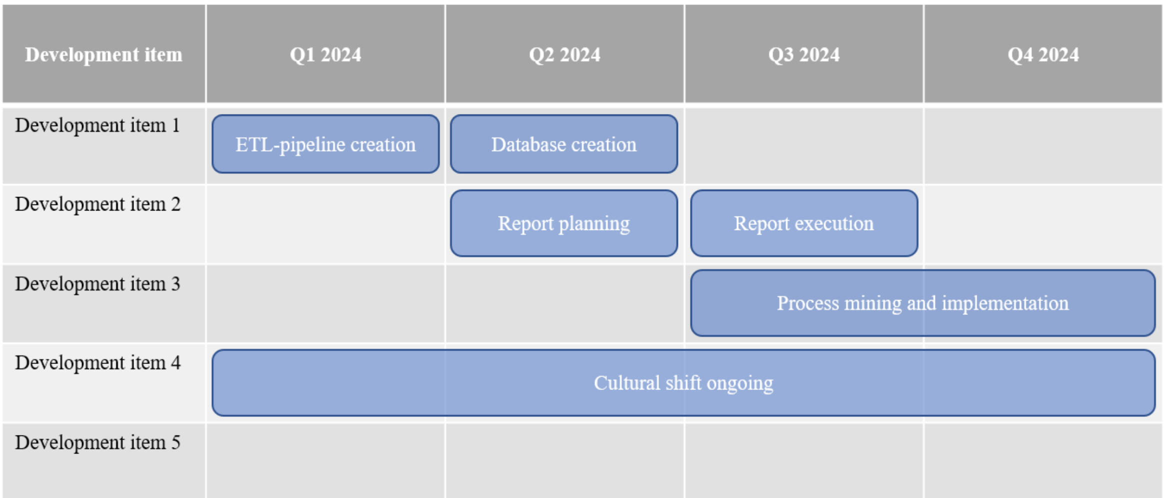

Figure 11. Case study Gant chart for development item implementation roadmap

As demonstrated in Figure 11, the development of the data-driven decision making in the case study organization should start in Q1 (Jan – Mar) 2024 by exploring the REST API:s provided by the CRM/DM system, sales system and the accounting system, and based on the findings, build the pipelines for extracting, transforming and loading the data into the data warehouse. After the exploration of the API: s and development of the pipelines, it is possible to see what kind of a data model, structure and tables should be deployed into the data warehouse. Hence, in Q2 (Apr – Jun) the data warehouse is deployed with a corresponding data model, and the data ingestion process is finalized (scheduling of the ETL-jobs and testing).

While the data warehouse is being built, the second development item, the dashboard creation, can be started in the Q2 as well. This will start in a form of drafting the needed visualizations, filters, drilling capabilities, timespans and a data model meeting all the above-mentioned criteria. The planning can be started right when the detailed content and form of data in the warehouse has been acquired. After planning, the report is built in the Q3 (Jul – Sep). After resources are released from the ETL work, the process development for the timer data generation can be started in the Q3. And for the whole timespan, the cultural shift from intuition-based decision making is taking place, enforced by the ongoing development of data-driven processes and decision making. At this stage, the market demand model is not yet moved forward.

4.4.5 Case study result summarization

As a conclusion, the results of the case study are the following:

- Method developed for selecting an MDDD model from the existing academic literature: comparison table (Table 2)
- Selected MDDD model: Maturity model of being a data-driven organization by Berndtsson, Forsberg, Stein and Svahn
- Maturity level of the case study organization: Level 2
- Main development items regarding MDDD for the case study organization: Figure 10
- Development roadmap for the case study organization: Figure 11

5 Conclusions

This thesis study aimed to define a level of maturity of data-driven decision making for a case study company, and while doing so, present a method for selecting an existing MDDD model from the academic literature, and assess how well the selected model can be applied outside of its original use case. In addition, the case study identified the main development items and an implementation roadmap for the case study organization, based on the observations of the maturity level measurement. This chapter presents the findings of the study, in addition to discussion on limitations of the study and future work.

5.1 Summary of findings

The maturity level of the case study company was identified to be 2 (on a scale from 1 to 4). This means, that the case study organization had already ongoing initiatives regarding data and analytics, along with data-driven decision-making processes, but the processes were still quite few, and the technical data architecture of the company did not allow full scale data-driven culture. No advanced analytics, data warehousing technologies, modelling or predictions were utilized. Five development items were identified, assessed and put on to a roadmap based on the maturity measurement discussions, from which the most significant value would come out by focusing on the data warehousing and reporting technology and processes of the company.

As a part of the case study, the thesis work presented a method for selecting the most optimal MDDD model from the existing literature. As there are lot of MDDD models available in the academic and non-academic context (Król, & Zdonek, 2020), and as it is a widely recognized issue that the models do lack standardization (Lasrado et al. 2015), it might be difficult for an organization to select a proper model for their need. To assist organizations and future research to identify and select MDDD models for their use, especially from the model usability point of view, the thesis study presented a 5-dimensional comparison table of MDDD models. In the comparison table, the assessed models are scored on different dimensions from 0 to 5, after which the model with the highest total score should be the most suitable for the user's needs.

The chosen MDDD model for the case study was a model developed by Berndtsson et al. (2015). The model utilized a framework of enabling factors for data-driven culture to offer guidelines on what organizational angles to consider in the measurement, and then a 5 dimension – 4 level matrix to assign the actual maturity level for an organization. As the actual maturity level matrices were very similar throughout different models (Król & Zdonek, 2020), the main advantage of the model by Berndtsson et al. was to offer the framework for studying the case study organization in a very convenient form. For example, models by Buitelaar (2018) and Eckerson (2009) provided long and specific questionnaires, which would have been difficult to fit and apply to the small and industrially differentiated organization, whereas Berndtsson provided a mind map shaped framework, which could be used in a more flexible way. And then on the other hand, some papers presenting MDDD models didn't offer any tools for gathering the information from the case study organization (e.g. LaValle et al., 2011).

As already expected, based on the literature review (Bititci et al., 2015; Cech et al., 2018), the main benefits for the case study organization of running through the process of measuring the maturity of its data-driven decision making, was not the defined level itself, but the findings made along the way, and the development items identified. There is a lot of room to grow in the future academic research regarding the utilization of the measurement results, as very few papers actually describe how the findings of the models were utilized by the organizations that used them. In the context of this thesis work, the identification of the main issues regarding the company's data-driven processes and their value/effort mapping was clearly the part that brought in the greatest added value.

The thesis work case study shows that existing models can and should be used in the research of measuring the maturity of data-driven decision making. With the use of the model by Berndtsson et al. (2015), there was no issues to define the MDDD level for the case study organization in an efficient way.

5.2 Limitations and future research

From the case study limitations point of view, it is important to notice that the case study organization was a relatively small size professional services company, and therefore the results of the study are applicable only in similar circumstances. For example, might be that

Berndtsson's model could have been too simple or high-level to capture the maturity of an organization, which would have had thousands of employees in several countries and business areas, or which would have been operating in a more data heavy industry. In addition, due to resourcing issues, the case study organization was not able to provide other personnel than the COO to comment or validate the discussion nor findings, which makes the results vulnerable for bias.

Considering the above-mentioned limitations, this study can be seen as high-level overview into the MDDD domain, in which one specific model is validated with a one specific organization. Hence, as part of the other below-mentioned future research needs around the area, repetitions for similar case studies are needed on other company sizes and industries. However, despite the limitations, hopefully this thesis study is able to make it easier for organizations and researchers to approach the topic, select existing maturity models from the literature based on their characteristics, and utilize e.g. the model by Berndtsson et al. (2018). This way more evidence and observations can be gathered on how to use and develop the models in the future.

By concentrating the future research more on existing work, the models can be developed even further, and what's even more important, they end up in usage more easily. Regarding the model by Berndtsson et al., it might be beneficial to finalize the model in a way, which would tie the enabling factors framework part even more closely to the maturity matrix part, so that the dimensions would be more aligned, and the user would always know, what discussion points in the enabling factors framework are affecting what dimensions and levels in the maturity matrix. In addition, more comprehensive documentation for the enabling factors framework usage would be beneficial (e.g. how to use the enabling factors framework in an interview to make sure everything relevant around data-driven processes has been discussed). Same goes for other future research as well: more concrete methodology descriptions should be developed and included to the MDDD publications, for making it easier for others to apply the models in practice. This would enable easier model validation with different industries and contexts, in addition to helping others build on existing models.

Another focus area for MDDD model future research would be to allow defining the scale of maturity on the fly while using the model. As for example, a multi-national globally listed industrial company has a very different viewpoint and possibilities regarding the maturity of data-driven decision making, compared to a small company operating in a not data-heavy industry. Even though the model by Berndtsson et al. (2018), enabled some interpretation and flexibility, the actual descriptions and limits between different levels were still quite strict. If the minimum and maximum maturity could be defined during the process of interviewing the case company, this would enable the model to consider the actual and realistic possibilities for the case study organization more easily.

References

Anderson, G. and Arsenault, N., 2005. *Fundamentals of educational research*. Routledge.

Ariyarathna, K. and Peter, S., 2019. Business analytics maturity models: a systematic review of literature. *Focus*, *3*(10), p.4.

Beck, S. E., & Manuel, K. (2008). Practical research methods for librarians and information professionals. New York, NY: Neal-Schuman.

Berndtsson, M., Forsberg, D., Stein, D., and Svahn, T., 2018. Becoming a Data-Driven Organisation. European Conference on Information System (ECIS).

Bititci, U., Garengo, P., Ates, A. and Nudurupati, S., 2015. Value of maturity models in performance measurement. International journal of production research, 53(10), pp.3062-3085.

Brooks, P., El-Gayar, O., and Sarnikar, S., 2013. Towards a Business Intelligence Maturity Model for Healthcare. IEEE, pp. 3807-3816.

Brynjolfsson, E., Hitt, L.M. and Kim, H., 2011. Strength in numbers: How does data-driven decisionmaking affect firm performance.

Bräutigam, D., Gerlach, S. and Miller, G., 2006. Business Intelligence Competency Centers. Hobo-ken: John Wiley & Sons, Inc.

Buitelaar, R., 2018. Building the data-driven organization: a maturity model and assessment. *Leiden University*.

Cardoso, E., Alcolea, J., Rieger, B., Schulze, S., Rivera, M. and Leone, A., 2013. Evaluation of the maturity level of BI initiatives in European Higher Education Institutions: Initial report from the BI Task Force @EUNIS. In Proceedings of the 19th International Conference of European University Information Systems (EUNIS 2013), Riga, Latvia, 12–14.

Cech, T.G., Spaulding, T. and Cazier, J., 2018. Data competence maturity: developing data-driven decision making. *Journal of Research in Innovative Teaching & Learning*, *11*(2), pp.139-158.

Chambers, R., 2002. Participatory workshops: a sourcebook of 21 sets of ideas and activities. London: Earthscan.

Chatterjee, S., Chaudhuri, R. and Vrontis, D., 2021. Does data-driven culture impact innovation and performance of a firm? An empirical examination. *Annals of Operations Research*, pp.1-26.

Davenport, T. and Harris, J. G., 2017. Competing on analytics: The new science of winning.

Dutta, A., Roy, R. and Seetharaman, P., 2022. An assimilation maturity model for IT governance and auditing. *Information & Management*, *59*(1), p.103569.

Eckerson, W., 2009. TDWI's Business Intelligence Maturity Model. The Data Warehousing Institute.

Elena, C., 2011. Business intelligence. *Journal of Knowledge Management, Economics and Information Technology*, *1*(2), pp.1-12.

Elsa, C., and Xiaomeng, S., 2022. Designing a Business Intelligence and Analytics Maturity Model for Higher Education: A Design Science Approach. Applied Sciences (12:4625), pp. 4625-4625.

Fraser, P., Moultrie, J. and Gregory, M., 2002. August. The use of maturity models/grids as a tool in assessing product development capability. In *IEEE international engineering management conference* (Vol. 1, pp. 244-249). IEEE.

Gibson, D.M. and Jefferson, R.N., 2006. "*The effect of perceived parental involvement and the use of growth-fostering relationships on self-concept in adolescents participating in GEAR UP*", Adolescence, Vol. 41 No. 161, pp. 111-125.

Goldenson, D. and Gibson, D.L., 2003. Demonstrating the Impact and Benefits of CMMI: An Update and Preliminary Results, Carnegie Mellon University, Software Engineering Institute.

Grimaldi, D., Diaz, J., Arboleda, H. and Fernandez, V., 2019. Data maturity analysis and business performance. A Colombian case study. *Heliyon*, *5*(8).

Grossman, R., 2018. A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, *38*(1), pp.45-51.

Gudfinnsson, K., Strand, M. and Berndtsson, M., 2015. Analyzing business intelligence maturity. *Journal of Decision Systems*, *24*(1), pp.37-54.

Gummesson, E., 2000. *Qualitative methods in management research*. Sage.

Gökalp, M.O., Gökalp, E., Kayabay, K., Koçyiğit, A. and Eren, P.E., 2022. The development of the data science capability maturity model: a survey-based research. *Online Information Review*, *46*(3), pp.547-567.

Halper, F. & Stodder, D., 2014. TDWI analytics maturity model guide. The Data Warehousing Institute.

Hamilton, B., 2015. The field guide to data science.

Heale, R. and Twycross, A., 2018. What is a case study?. *Evidence-based nursing*, *21*(1), pp.7-8.

Hostmann, B. & Hagerty, J., 2010. IT-score overview for business intelligence and performance management.

HP, 2007. The HP Business Intelligence Maturity Model: describing the BI journey.

IBM, 2009. Overview MM. Finding Your Level of BI and Performance Management Maturity: BI and Planning for Midsize companies Introducing IBM cognos. IBM.

Johnson, D., 1995. Research methods in educational management.

Kim, S. and Ji, Y., 2018. Gap analysis. *The international encyclopedia of strategic communication*, pp.1-6.

Król, K. and Zdonek, D., 2020. Analytics maturity models: An overview. *Information*, *11*(3), p.142.

Lahrmann, G. and Marx, F., 2010. Systematization of maturity model extensions. In Global Perspectives on Design Science Research: 5th International Conference, DESRIST 2010, St. Gallen, Switzerland, June 4-5, 2010. Proceedings. 5 (pp. 522-525). Springer Berlin Heidelberg.

Lasrado, L.A., Vatrapu, R. and Andersen, K.N., 2015. Maturity models development in is research: a literature review.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., and Kruschwitz, N., 2011. Big Data, Analytics and the Path from Insights to Value. MIT Sloan Management Review (52:2), pp. 21-32.

Logi Analytics, 2019. The 5 Levels of Analytics Maturity: From Basic BI to Sophisticated Differentiators; ebook; Logi Analytics: McLean, VA, USA. Available online: https://goo.gl/8x7vgQ (accessed on 28 September 2019)

Microsoft, 2023. Introduction to Dataflows in Power BI. Available at: https://learn.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-introduction-self-service (Accessed: 18 November 2023).

Miller, G.J., Bräutigam, D. and Gerlach, S.V., 2006. *Business Intelligence Competency Centers: A team approach to maximizing competitive advantage*. John Wiley & Sons.

Müllert, N. and Jungk, R., 1987. Future workshops: how to create desirable futures. London: Institute for Social Inventions.

Nott, C., 2014. Big data & analytics maturity model — IBM big data & analytics hub.// https://www.ibm.com/analytics/big-data-analytics.

Ponniah, P., 2004. Data warehousing fundamentals: a comprehensive guide for IT professionals. John Wiley & Sons.

Provost, F. and Fawcett, T., 2013. Data science and its relationship to big data and data-driven decision making. *Big data*, *1*(1), pp.51-59.

Rigby, S. & Contreras, G., 2015. Adobe analytics operational readiness playbook.

Runkler, T.A., 2016. *Data analytics*. Wiesbaden: Springer Fachmedien Wiesbaden.

Schmarzo, B., 2016. Big Data MBA: Driving Business Strategies with Data Science.

Shaaban, E., Helmy, Y., Khedr, A., & Nasr, M., 2011. Business intelligence maturity models: Toward new integrated model. In The International Arab Conference on Information Technology (pp. 11-14).

Simpson, J.A., Weiner, E.S.C., 1989. The Oxford English Dictionary, 2 ed., Oxford, UK: Oxford University Press.

Soneryd, L. and Amelung, N., 2016. Translating participation: Scenario workshops and citizens' juries across situations and contexts. In *Knowing governance: The epistemic construction of political order* (pp. 155-174). London: Palgrave Macmillan UK.

Wilson, V., 2012. Research methods: interviews.

Yin, R.K., 2009. *Case study research: Design and methods* (Vol. 5). sage.

Ørngreen, R. and Levinsen, K.T., 2017. Workshops as a research methodology. *Electronic Journal of E-learning*, *15*(1), pp.70-81.

Öberg, J. and Hernwall, P.A.T.R.I.K., 2016. Participatory design with teachers: designing the workshops. In *Designs for Learning, Proceedings of the 5th International Conference on Designs for Learning* (pp. 269-282).

Appendix 1. Reference list of the comparison table (Table 1) by Król and Zdonek

1. Grossman, R.L., 2018. A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, *38*(1), pp.45-51.
2. Piyanka, J., 2019. The Analytics Maturity Quotient Framework. *Aryng LLC. Available online: https://goo. gl/3RwiYJ (accessed on 28 September 2019)*.
3. Blast. Analytics Maturity Assessment. Blast Analytics & Marketing. Available online: https://www.blastanalytics.com/analytics-maturity-assessment (accessed on 28 January 2019)
4. 5 Areas to Assess Using the DAMM—Data Analytics Maturity Model. Association Analytics. Available online: https://associationanalytics.com/blog/5-areas-assess-data-analytics-maturity-model/ (accessed on 28 September 2019).
5. Davenport, T.H., Harris, J.G. and Morison, R., 2010. *Analytics at work: Smarter decisions, better results*. Harvard Business Press.
6. Gartner Survey Shows Organizations Are Slow to Advance in Data and Analytics. Gartner Newsroom. Available online: https://www.gartner.com/en/newsroom/press-releases/2018-02-05-gartner-survey-shows-organizations-are-slow-to-advance-in-data-and-analytics (accessed on 28 January 2019).
7. Logi Analytics. The 5 Levels of Analytics Maturity: From Basic BI to Sophisticated Differentiators; ebook; Logi Analytics: McLean, VA, USA. Available online: https://goo.gl/8x7vgQ (accessed on 28 September 2019).
8. Cardinal Path. Benchmarking Your Organization's Analytics Maturity? ebook; Cardinal Path: New York, NY, USA. Available online: https://goo.gl/KgeAwH (accessed on 28 September 2019).
9. Five Steps to Analytical Maturity. A Guide for Pharma Commercial Operations, White Paper; SAS & PharmaVOICE: Stockholm, Sweden, 2014
10. Halper, F.; Stodder, D. TDWI Analytics Maturity Model Guide. Interpreting Your Assessment Score; TDWI Research; The Data Warehousing Institute: Renton, WA, USA, 2014.
11. Hamel, S. The Web Analytics Maturity Model. A Strategic Approach Based on Business Maturity and Critical Success Factors. Available online: https://goo.gl/96nPyr (accessed on 28 January 2019).