DATA-DRIVEN DECISION-MAKING MATURITY

Herjan Nijzink
University of Twente
PO Box 217, 7500 AE Enschede
the Netherlands
h.j.nijzink@student.utwente.nl

ABSTRACT

The consensus in literature is that data-driven is changing the decision-making process in a positive way. However, businesses often do not know where to start when introducing this new concept. This paper solves that problem by creating a simple, yet effective measurement instrument that not only identifies your current data-driven decision-making maturity level but also provides guidelines on how to improve. Furthermore, it stimulates continuous improvement to break down unattainable goals into small, manageable steps. A case study was performed to validate the model in practice. The results show positive signs that the model is working according to its purpose. However, to prove validity, more research is needed.

Keywords

Data-driven, Decision-making, Maturity, Measurement Instrument, Continuous Improvement

1. INTRODUCTION

Every single day, companies face decisions. Some decisions are really important and affect the future of the company whereas other decisions are of less importance. All decisions are taken based on a knowledge source. According to literature, two kinds exist: Data and Experience (Divan, 2018). Data-driven decision-making (also referred to as DDDM) uses data as a basis for the decision. For example, data has been collected about sales, and it shows that a certain product has been exceptionally popular for the past weeks. Based on this knowledge, a manager can make decisions regarding the production of this product. Experienced-based decision-making uses the experience that a manager has with similar decisions that he/she has taken in the past. If that decision turned out well in the past, why not take the same decision now? I like to add another knowledge source which is only slightly touched upon in literature, namely theory. Thiess et al. (2018) mention the concept of human judgment, which can either refer to your own knowledge or to knowledge from other human beings. Theory-based decision-making uses the knowledge (and experience) of other people. Most likely, researchers have already studied these kinds of situations to come up with a theory about what to do when facing such a decision.

Even though several separate categories of decision-making can be identified, it does not mean that a decision consists of only one of them. A decision is often based on both data and experience (Provost et al., 2013). Although data-driven decision-making is becoming more and more popular, not all decisions can be made solely on data. If for example the data shows that 10 new branches should be built in several countries, no manager will blindly follow this advice and decide on immediately building 10 new locations just because the data tells them to do so. Reason is that this would come with such high risk that common sense tells us this might not be smart. First, we need to do in-depth research to identify potential consequences. Even though it is useful that the data tells us to expand, experience (or else theory) tells us to do so step by step. This means that for this type of decision, data can provide advice but is not leading.

The adoption of data-driven provides more objectivity and reveals insights that humans would never think of (Vohra, 2016; Streifer, 2004). Measuring the extend in which a company makes use of data is called its level of maturity. Measuring this level of maturity requires a measurement instrument like a maturity model. Such models should indicate the current maturity level as well as identify guidelines for improvements (Selladurai et al., 2020).

However, current literature does not provide a data-driven decision-making maturity model. Therefore, the goal of this paper is to construct such a maturity model to help businesses improve their DDDM processes. Subsequently, the model will be applied to a company called Coulisse during a case study. This is needed to validate the model in practice. Coulisse has about 180 employees and is a global specialist in window coverings: it designs, produces, and sells window decorations. This company owns a lot of data regarding products, customers, employees, and sales. However, they currently make little use of it. According to their information manager, it is only now that they start to see the potential value of these data. Therefore, they first want to research the best way in which the company can take advantage of the data.

An important objective of this paper will be to assess what capabilities are required to be able to transition into a mature data-driven company. Once these capabilities are identified, they will serve as input for a data-driven decision-making maturity model.

The main goal of this paper is to come up with a measurement instrument for identifying the current maturity level of data-driven decision-making and to provide guidelines on how to reach a higher maturity level.

Last objective of this project is to validate the data-driven decision-making maturity model in practice. How should a company apply the model, and how can the model help them to improve their data-driven decision-making? Questions like these will be answered by a case study at Coulisse.

Based on these objectives, we have the following research questions:

MRQ

What is a valid maturity model to measure and improve the level of data-driven decision-making?

SQ1

How can a company become more mature regarding data-driven decision-making?

The word 'valid' has different definitions depending on the context, and therefore requires some explanation. For this paper, it regards a maturity model which means we can make use of the second definition stated at Merriam-Webster (n.d.-b): well-grounded or justifiable, being at once relevant and meaningful. Well-grounded and justifiable regards the design of the model and the components used. These are almost all based on the literature (a few are based mostly on practice). Whether the model is being relevant and meaningful is determined during the case study when the model will be applied to practice.

The model should be proven valid when applying it to practice. This means that all components need to be grounded or justifiable, and also relevant and meaningful. To prove full validity of the model, a lot of research is needed. To show that all components are relevant, the model will need to identify challenges on multiple levels. Because if (according to the model) processes always get stuck on the same level, this would indicate not all of its components are relevant or meaningful.

Furthermore, the expectation is that the DDDM maturity model can be used in all companies, not depending on the type of company or the sector they are in. It should be easy generalizable and therefore applicable in every context.

2. METHODOLOGY

An important goal of this paper is to construct a measurement instrument. Since this can be regarded as designing an artifact, it suits best to make use of the Design Science Research Methodology (also referred to as DSRM). This is the main method throughout this project. The DSRM (see figure 1) clearly identifies separate phases which follow a specific sequence. However, as can be seen in figure 1, arrows starting at the evaluation phase point back towards earlier phases. This means that whilst evaluating, new insights and ideas might come up that require changes to be made to work that was done during previous phases.

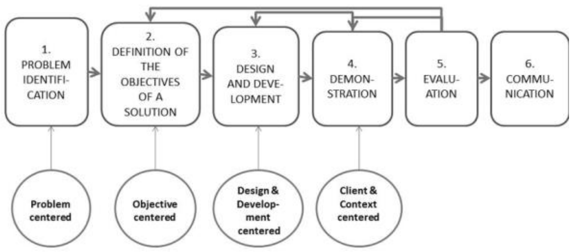

Figure 1: Design Science Research Methodology Model (retrieved from: Pulkkinen, 2013)

Another goal of this paper is to perform a case study which serves as a validation method for the measurement instrument. For this part, the Action Research Methodology (also referred to as ARM) is used (see figure 2). This is a well-known methodology when performing a case study which has been applied in lots of research already. According to D.R. Corey, "Action Research is the process by which practitioners attempt to study their problems scientifically in order to guide, correct and evaluate their decisions and actions". This method aims to develop scientific knowledge while acting to solve real problems at the same time. The action research model consists of four sequential phases which are part of a loop that can be executed multiple times.

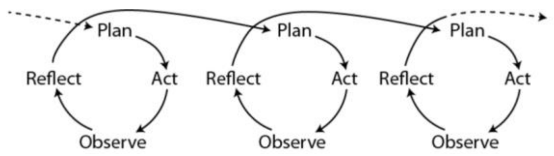

Figure 2: Action Research Cycles (retrieved from: Ejbye-Ernst & Jorring, 2017)

Often, when both the DSRM and the ARM are being used at the same time, these are combined into the Action Design Research Methodology (also referred to as ADRM). The ADRM (see figure 3) is proposed by Sein et al., (2011) to "conduct Design Research that recognizes that the artifact emerges from interaction with the organizational context even when its initial design is guided by the researchers' intent. This research method generates prescriptive design knowledge through building and evaluating ensemble IT artifacts in an organizational setting" (Sein et al., 2011, p.40). The dual mission of this method is to both make theoretical contributions as well as assist in solving current problems of practitioners (Benbasat & Zmud, 1999; Iivari, 2003; Rosemann & Vessey, 2008 as referenced in Sein et al., 2011, p. 38). Whereas traditional DSR methods often encounter a disconnect between the development of artifacts and their application in organizations, the ADR method takes into account the role of organizational context during the design and deployment of the artifact (Thiess et al., 2018). Since these definitions apply to this project as well, the decision was made to make use of the ADRM instead of two different methods (DSRM & ARM).

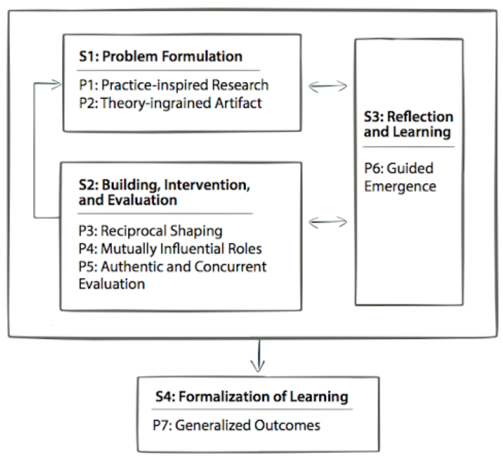

Figure 3: Action Design Research Model (retrieved from: Sein et al., 2011)

The first phase consists of formulating the problem. This can be any kind of issue perceived in practice or anticipated by researchers. Once a problem has been identified, its scope is being determined after which the research questions can be formulated. The second phase continues on the results from the first phase. Now the initial design of the artifact can be generated, which is further shaped once more relevant information becomes available. This second phase is iterative, meaning that the building, further shaping and evaluation are interweaved. Outcome of this phase is the realized design of the artifact. Phase three parallels the first two phases. It is no longer just about offering a solution to one specific problem, but rather about applying the results to a broader class of problems. The objective of the fourth and final phase of ADRM is to formalize the learnings. Goal is to generalize the outcomes of the project into general concepts. For more detailed explanations of the ADRM, please see Sein et al., (2011). To provide some more explanation as of how the ADRM has been applied to this research, figure 4 is included.

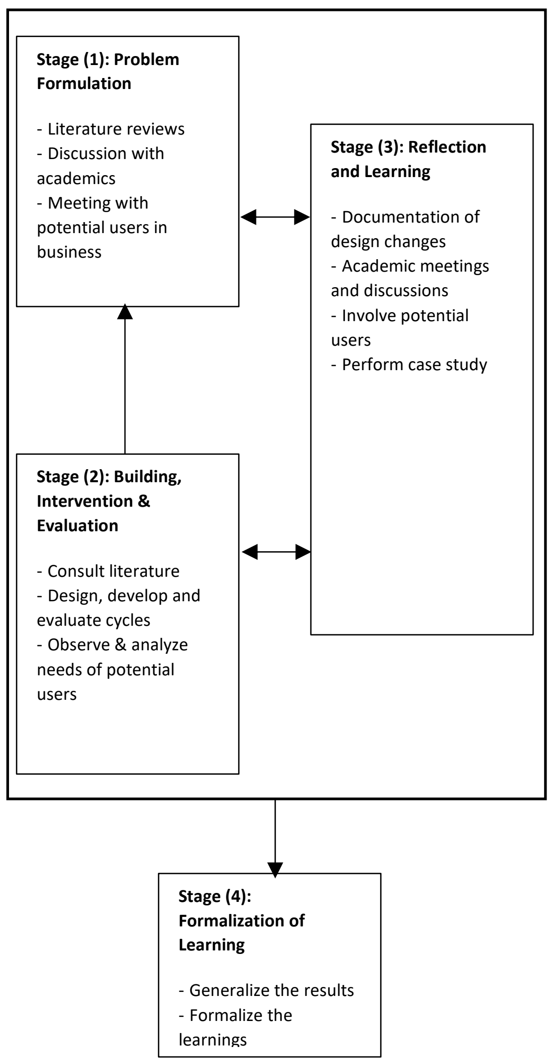

Figure 4: Action Design Research applied to this research

Regarding the design of a maturity model, no shared vision exists on which approach should be followed (Mettler & Rohner, 2009; Pöppelbuß & Röglinger, 2011). For this project, two main methods were used. First, six generic phases proposed by de Bruin et al., (2005) were explained and applied to the construction of the measurement instrument. Even though not all phases are part of the scope of this project, the relevant ones are described in section 4.3. Furthermore, guidelines proposed by Pöppelbuß & Röglinger (2011) were used. These guidelines exist of design principles and are useful for the design of constructing the measurement instrument. More details about these design principles can also be found in section 4.3.

3. State of the Art

What is meant by data-driven and decision-making? The concept of decision-making has been around for a long time whereas the concept of data-driven is relatively new. Important papers that set the foundations for later work about decision-making date from decades back (Edwards, 1954; Mintzberg et al., 1976). Articles about data-driven on the other hand are more recent (Khan, 2019).

3.1 Decision-making

Edwards (1954) discusses the traditional way of decision-making when speaking about the 'economic man': the theory of riskless decision-making. This means that when making a decision, all options and their related outcomes are fully known, and thus a rational decision is made. The other option would be risky decision-making, in which the outcomes are not (fully) known. Herbert Simon had a problem with the traditional decision-making process, since he argued that decision makers not always optimize their decision-making process. Reason is that information and/or computation was required that no human being could possess. Therefore, he introduced 'bounded rationality' and 'satisficing'. Satisficing means a good enough solution, which not necessarily equals the most optimal solution. This is useful when incomplete information or limited computation is available (Herbert, 1955; Herbert, 1956; Brownlee, 2007).

According to Klein (1995), the traditional view is focusing only on part of the process, called the decision event. This means that the decision taker identifies the alternatives, weighs the consequences, and does the decision-making. Recent articles regarding decision-making still recommend this traditional view, even though some changes are made (Gray, n.d.). Klein offers an alternative way to make decisions. That is to define the situation and based on similar experiences selecting the best action. This action is then evaluated by identifying potential consequences and undesirable effects. If none are found, the action can be implemented.

Eisenhardt et al. (1992) discuss the paradigms of 'rationality vs bounded rationality', 'politics and power' and 'garbage can'. They state that decisions are no longer made fully rational and that executives are willing to put aside personal disagreements way more often than the traditional view was suggesting.

Dean et al. (1996) question whether procedural rationality and political behavior influence decision success. By a regression analysis, they concluded that different decision-making processes result in different outcomes, instead of every process having the same result. This makes that different decisions require different decision processes.

It might seem that all these decision processes are complicated causing you to lose a clear overview. However, decades ago, Mintzberg et al. (1976) already explained that even though a decision process might seem unstructured at first, all these processes can be modelled in a basic structure. They constructed a model that fit all 25 decision processes that were studied during their research. This suggests that all existing decision processes are structured in some way. And even if some decision problems turn out to be weakly structured, Decision Support Systems could offer a helping hand according to Althuizen et al. (2012). These systems can support the decision-maker in complex decision processes by means of analyzing a huge amount of data. However, the problem is a gap between the evaluation of the users and the actual performance of the support systems. Even if the systems performed well, more than half of the time the users fail to acknowledge this. This means that they write a negative evaluation whereas the system had a positive impact. And because of these negative evaluations, potential users are less likely to start using such a system. Only once this technology gets more adopted, the intended users will start to see its potential value.

3.2 Data-driven

Being data-driven means that data is used to come to a conclusion, instead of relying on your intuition and experience (Thiess et al., 2018; Vohra, 2016; Brynjolfsson et al., 2011; Divan, 2018). However, sometimes problems also require a combination of data-driven and experience (Provost et al., 2013).

Although data-driven offers a lot of advantages, not every business is ready to implement it. According to Mcelheran & Brynjolfsson (2017), data-driven gets more often adopted if the company has high levels of IT (resources), educated workers (skills), greater size (larger company) and better awareness. This is in line with Berg et al. (2018) who state that especially startups might not be able to correctly implement a data-driven culture, resulting in the inability to gain its value. Causes are:

- Lack of sufficient data
- Use of outdated data
- Lack of skills and/or resources

First reason is that startups often encounter limitations regarding their data in terms of volume, velocity, and variety (Berg et al., 2018). Even though they might be collecting data already, no data from previous years is available for processes like data mining and analytics. Also, data from a few months ago might already be outdated and irrelevant since the company changes so quickly in its early stages (Vohra, 2016). Furthermore, they might not have the skills and/or resources (like time and money) to transition into a data-driven culture. Data needs to be complete; data management should be correct and data analytics skills and knowledge must be in place to make use of data products (Khan, 2019). Also, startups often gain their data from one prototype with a few users, which makes it hard to generate the value. Therefore, young companies have no real choice than to trust on the experience of its employees and especially its board of directors instead of trying to be data-driven too soon.

Larger, more mature companies on the other hand will not encounter these issues. However, it often proves to be hard for those companies to adopt new or unfamiliar strategies.

One of the largest problems for companies that make the transition to data-driven is the lack of having a fixed end goal before collecting and analyzing the data. There are so many possibilities out there nowadays that it is more important than ever to set specific goals. Lots of young businesses spend way too much time mining the data without getting useful insights (Vohra, 2016). Therefore, you should always have a clear analytical objective. For example, if you want to assess an opportunity, different ways of analytics should be used compared to when diagnosing a business problem. The clearer the end goal is, the more focused and valuable the analysis will be.

3.3 Data-driven decision-making

Data-driven decision-making is about decisions that are made based on data. This is different than the traditional decision-making which mainly uses previous experiences to make a decision. DDDM uses data science, data processing and data engineering to come to a decision (Thiess et al., 2018; Provost et al., 2013).

An increasing amount of literature is suggesting that DDDM generates business value, especially from an economic perspective. Davenport & Harris (2007), for example, found a positive correlation between the adoption of analytics in organizations and their annual growth rates (based on a survey amongst 32 companies). Brynjolfsson et al. (2011) examined by a survey among 179 firms that adopting DDDM increases output and productivity by 5%-6% higher than expected when considering their other investments and information technology usage. Also, a recent study by Müller et al., (2018) showed that big data and analytics on average increases firm productivity of 4%, with some industries reaching even more than 7%. They did so by examining more than 800 firms over a period of seven years. A similar outcome was stated by Wu et al. (2016). However, another important finding of them was that most of the value of DDDM is in the enabling of continuous improvements; something which is perfectly aligned with the results of my own research. These findings are in line with many qualitative studies that also report positive relations between implementing DDDM and the business value (e.g. Manyika et al., 2011; vom Brocke et al., 2014; Someh & Shanks, 2015; Côrte-Real et al., 2017).

To turn data into value, one should consider behavioral aspects of human decision-making (Kahneman, 2003; Thaler, 1980; Tversky & Kahneman, 1992). Human decision-making might be affected by cognitive biases since humans tend to apply simplified heuristics when making (highly uncertain) decisions because of their limited information processing capacities (Tversky & Kahneman, 1974).

Really important in DDDM are the processes prior to the decision-making. If these processes lack quality, it drastically affects the quality of the decision. Most important factor is the quality of the data used for analysis (Moreno, 2017). But DDDM also requires skills and technology like business intelligence, data mining and analytics for the actual analysis of the data. All these components need to be in place to be able to use the DDDM concept

(Divan, 2018; Thiess et al., 2018; Provost et al., 2013). Each step of data collection and management must lead toward acquiring the right data and analyzing it to get the actionable intelligence necessary for making data-driven business decisions. Based on this knowledge, the causal model in section 4.4 has been created.

Before we end this section about DDDM, let us first look for some more information about the changes that adding data-driven to the decision-making process causes. What does it mean if the traditional decision-making process is transitioning into a more data-driven process? What changes does this result in? By analyzing the literature, the final part of this section tries to come up with a clear answer that explains why certain aspects of the decision-making process will undergo a drastic change.

The consensus in literature is that companies benefit using data-driven decision-making. Several general aspects are mentioned in different papers as of how the concept of data-driven does impact the decision-making process. This section will list and explain these aspects. However, before we get to them, lets first make clear what exactly is meant by the word ‘impact’.

The term ‘impact’ is used in a lot of different ways which can be confusing. Out of all definitions out there, two main interpretations are identified: ‘the action of one object coming forcibly into contact with another’ and ‘a marked effect or influence’. The important difference is that the former looks for causes of an effect whereas the latter looks for effects of a cause (Hearn, 2020). For this case, we will use the first definition since we are looking for the difference of a predefined indicator (the decision-making process) with the intervention of data-driven and without the intervention of data-driven.

When integrating data-driven into the decision-making process, the first thing one must realize is that decisions are no longer made according to traditional decision making. Now this might sound obvious, but it is an important distinction that is made here. Traditional decision making mainly makes use of previous experiences. These experiences are then used to make a similar decision. However, when adding data-driven to the process, the importance of experience drastically declines and is replaced by data.

Furthermore, one of the most fundamental changes that the transition to more data-driven is causing might be the focus on the underlying process. No longer is it just about the decision-making itself, but more and more of the attention is shifting to the steps prior to that. When deciding based on data, first you need to be sure that the data is correct and reliable. High data quality reduces the chance of making wrong decisions (Kleindienst, 2017). However, securing data quality is not enough since the data needs to be processed in the right way as well. Also, skills and technology like business intelligence and data analytics are required to use data for making actual decisions (Divan, 2018; Thiess et al., 2018; Provost et al., 2013).

Also, the decisions are more objective and therefore accurate. Objectiveness is obtained by using facts that are extracted from the data. This is different than the traditional way of decision making, which mainly makes use of opinions by means of personal experiences. To understand how a higher rate of accuracy is achieved, it is important to consider the behavioral aspects of human decision-making. As mentioned above, when humans have to make complex decisions, they tend to simplify the situation because of their limited information processing capacity (Tversky & Kahneman, 1974). This results in a lower accuracy compared to for example algorithms (Grove et al., 2000).

Another change is the discovery of new insights. Humans can determine a limited number of insights because of their restricted capabilities. However, data (and especially data analysis) provides many new insights that humans never have thought of (Vohra, 2016; Khan, 2019; Streifer, 2004). The result is that without data analysis, important information is missed out on when coming to a decision.

Last, and this is the biggest challenge, it is important to determine the right tradeoffs for a decision. Can the decision be made solely based on data, or is there a need for other sources as well? Using just data might sound attractive but might not always be the best way to go since often also human aspects need to be included. Therefore, some problems require a combination of data-driven and experience (Provost et al., 2013).

Even though lots of literature can be found regarding data-driven decision-making, an important gap was identified during the literature study. Thiess et al. (2018) state that although lots of papers mention the value of DDDM, there is a lack of knowledge on how to successfully employ DDDM in organizations. This paper provides the first step in filling this gap.

Let us end this literature part about DDDM with a quote from James Barksdale, former CEO of Netscape. His words show that the preference of the decision maker (one of the variables in the measurement instrument) is still important in the decision-making process, when he famously said:

"If we have data, let us look at data. If all we have are opinions, let us go with mine".

4. DESIGN OF A DDDM MATURITY MODEL

4.1 General info

Merriam-Webster defines maturity as "the quality or state of being mature" (Merriam-Webster, n.d.-a). This implies that the goal is to reach the state in which the maturity is highest, meaning that to do so several other levels need to be completed first. Smits et al. conclude that maturity models are therefore helpful in finding better solutions for change since weak spots are identified and guidelines for improvement might be provided by the model. (Smits & Hillegersberg, 2015).

Before maturity models were introduced in literature, lots of well-known models already existed that were based on a staged sequence of levels. Examples include Maslow's hierarchy of human needs (1954), Kuznets' economic growth model (1965) and Nolan's progression of IT in organizations (1970s). These can be regarded as forerunners of the many maturity models that would be introduced in the years to come.

Arguably the most well-known maturity model was published in 1991. This model was requested by the U.S. Department of Defense and was developed by the Software Engineering Institute. Reason for this request was that many US military projects involving software subcontractors ran over-budget and were not completed on schedule (CIO wiki, n.d.). To find the cause of the problem, the Capability Maturity Model (also referred to as CMM) was developed, and version 1.0 was published in 1991 (Paulk et al., 1991; Paulk, 2009). The Capability Maturity Model provides a framework for organizing evolutionary steps into five maturity levels that lay successive foundations for continuous process improvement (just like the model created in this paper). The CMM is widely used and serves as the foundation of many other maturity models. Since its first version, the CMM and its many extensions have been used by organizations worldwide as a general and powerful instrument for understanding and subsequently improving general business performance (CIO wiki, n.d.).

Due to the CMM, the origin of maturity models lies in the domain of software engineering. However, since the launch of CMM, the application of maturity models has been increasing a lot. Hundreds of maturity models have been constructed and are applied in over 20 domains (Wendler, 2012). Maturity models nowadays are no longer connected to certain application domains, but rather refer to key managerial dimensions such as people, processes, and organizational capabilities (Navarro, 2014). A well-known example is the organizational growth model by Greiner (back in 1972), in which he describes five stages of growth of a company. As the age of an organization matures, its size will grow. Each growth phase leads to a so called 'crisis', which represents a management problem that needs to be solved before further growth is possible (Greiner, 1998).

4.2 Purpose

Maturity models assess organizational capabilities of a specific domain in a stage-by-stage manner based on a set of criteria (Pöppelbuß et al., 2011; de Bruin et al., 2005). They offer organizations an effective opportunity to measure the quality of their processes. Other papers add to this definition when stating that 'maturity models are a business instrument that facilitates change or improvement by providing a framework based on performance parameters which both assess the current organizational capabilities and provide a path for improvement' (Gregory & Roberts, 2020). Continuous improvement is an important aspect of maturity measurement. If a high maturity level is established, chances increase that errors made during the process will lead to improvements in quality or use of resources. When having a low maturity level, these issues tend to get solved less often. Although aspects of maturity models are discussed in several papers, the papers of Pöppelbuß et al. and de Bruin et al. are fully focusing on the purpose of maturity models and their construction. Therefore, these papers will serve as most important knowledge sources when constructing our own maturity model.

Even though not all papers use the exact same definition, everyone agrees that maturity models should be regarded as measurement instruments. These instruments are used to measure the status of a company regarding a specific domain. When applying this to DDDM, we want to measure the maturity level of DDDM in a company.

Pöppelbuß et al. (2011) suggest in their paper that the purpose of a maturity model can be threefold: descriptive, prescriptive, and comparative. A model is purely descriptive when the current situation is assessed. This is useful if you want to get an idea about how a certain domain inside your company is performing. A prescriptive model on the other hand provides recommendations on how to both find and reach desirable maturity levels to increase business value. These prescriptive models are also used in a descriptive way, since first one needs to gain insight in the current situation before one can start to identify a desirable level. Once the current level is identified, the prescriptive part comes in by recommending guidelines to reach the desired level. Comparative models allow for internal or external benchmarking. They can be used to either compare the maturity of a specific domain to another internal domain, or to compare across industries or regions. Furthermore, maturity models exist that combine all three: holistic models. These integrate the descriptive, prescriptive, and comparative parts (Navarro, 2014).

Even though the discussed types of maturity models are distinct, some argue that they all represent part of a maturity model lifecycle. Some are only describing the as-is situation. As mentioned above, these can be evolved to a prescriptive model when adding guidelines on how to reach a desired maturity level. Once the model has been applied in multiple organizations, enough data is available to start using it comparatively (de Bruin et al., 2005).

This project requires a prescriptive model. The company that takes part in the case study wants to know how to reach a higher DDDM maturity level. This indicates that there is no need to compare their current situation to other companies or to another department inside their own company. Furthermore, a descriptive maturity model would only provide an assessment of their current situation. And although that is an important first step, it is not enough. Once the situation is assessed, guidelines should be provided on what is required to reach a higher maturity level.

4.3 How to construct

Several papers create their own maturity model and describe how to do so. De Bruin et al. (2005) propose and explain six generic phases that should be used when creating a maturity model: scope, design, populate, test, deploy and maintain (table 1).

PHASE	EXPLANATION
Scope	Identify which domain is targeted
Design	Design model to needs of intended audience
Populate	What needs to be measured and how?
Test	Test for validity, reliability, and generalizability
Deploy	Make the model available for use
Maintain	Maintain the model once more info is available

Table 1: Six generic phases for creating a maturity model (retrieved from: De Bruin et al., 2005)

Scoping has already taken place since the topic of this research is DDDM. The design phase needs to consider the intended audience. This is taken care of by talking to Coulisse (the company that takes part in the case study) during this phase already to identify their needs. Populating the maturity model will be done based on the causal model of section 4.4 which contains the concepts that need to be measured. The measurement will be done during the case study by means of observations and interviews. Testing the maturity model will be done by applying the model to two processes at Coulisse. However, since Coulisse is only the first company that uses the model, validity cannot be fully proven. Therefore, the testing phase will not be finished during this project. Deployment will only be done by means of publishing this paper online. The maturity model that has been constructed during this paper will not be made available separately. The last phase, maintaining the model, is not included in the scope for this project, and will therefore not be executed.

As mentioned above, the paper of Pöppelbuß et al. (2011) provides guidelines on how to construct maturity models. Their first research question is looking to identify design principles for maturity models to make them useful in their domain and fulfill their purpose. The design principles are divided into three categories: descriptive, prescriptive, and comparative (see table 6 in appendix B). Since comparative models depend for a large part on external factors, they decided to not include design principles for them. This leaves three categories: basic principles (which relate to both descriptive and prescriptive models), descriptive principles and prescriptive principles. Descriptive models should include their own principles as well as the basic principles, whereas prescriptive models should include all three categories. An important note is made in the paper of Pöppelbuß et al. that not every maturity model necessarily must include each design principle. The list rather serves as a guideline and checklist.

Since this project constructs a prescriptive model, all design principles are relevant and should therefore be taken into consideration. However, not all of them must necessarily be implemented at all cost.

Next to the six phases of De Bruin et al. and the design principles of Pöppelbuß et al., we find that several more guidelines are provided in the literature. Even though some are specific to certain types of maturity models, others are generic and therefore applicable to all maturity models. Especially these general guidelines could be useful, so they should not be neglected when developing a data-driven decision-making maturity model from scratch.

One of these guidelines is to make use of multiple levels. According to several papers, based on an assumption of predictable patterns of change and evolution, maturity models usually include a sequence of levels that together form an anticipated, desired, or logical path from an initial state to maturity (Becker et al. 2009; Gottschalk 2009; Kazanjian and Drazin 1989). De Bruin et al. (2005) add that the most used way of measuring the level of maturity is by means of a five-point Likert scale with '5' being the highest maturity level possible. This information tells us that we should make use of multiple levels, preferably somewhere around five.

Another guideline is about how to implement those levels. Literature suggests there are two approaches for implementing a maturity model. One is a top-down approach proposed by Becker et al. (2009), which means that first a fixed number of levels is decided upon before confirming this number with characteristics/variables that support the initial assumptions. The other is a bottom-up approach suggested by Lahrmann et al. (2011). The bottom-up approach first comes up with the characteristics/variables before they are being assigned to maturity levels. This project uses the bottom-up approach since the characteristics/variables have already been determined beforehand by means of the causal model (the process of determining the characteristics and variables of the model is described in more detail in sections 4.4, 4.5 and 5.1). Only once the variables are known, the number of maturity levels is determined.

4.4 Causal model

Since this research makes use of the bottom-up approach, the first step is to come up with the variables before we use the other guidelines from section 4.3 to design the maturity model itself. The variables that will be part of the maturity model are based on the so called 'causal model'. Before constructing the causal model, it is important and helpful to first identify what components should be part of the model. As mentioned above, the components are divided into dimensions and variables. Dimensions are the main components of the causal model. They are identified by means of analyzing the literature, for example what has been described in section 3.3. The variables are influencing the dimensions and are identified by consulting literature.

After having studied the literature regarding DDDM, the causal model in figure 5 was constructed. It visualizes the concepts and processes on which DDDM is based. The blue rectangles represent the dimensions of the decision-making process. The green boxes represent the variables that influence these dimensions. Every arrow that connects two dimensions, visualizes a transition, and contains risk factors.

The starting point of the whole process prior to the decision making is the reality. This dimension is in place at every company, even though the situation might differ. Analyzing the current situation leads to new insights. Reality therefore provides lots of opportunities and is the start of any new idea or process. Reality offers a chance to collect data by means of measurement methods.

But immediately questions arise regarding the reliability of the data: is the data a correct representation of reality? Also, is the right data being collected or is any useful data missing? This indicates that three variables are important for the data dimension: Data quality, Data collection procedures and Data volume (Moreno, 2017; Kleindienst, 2017).

Once the data is collected, it is important to organize it (Demchenko, n.d.). The data should be organized and combined in a clear way, preferably making use of one central data warehouse/lake instead of multiple data sources. This way of managing data is important to be able to use the data in an effective way in the next dimensions. Data management is the process of ingesting, storing, organizing, and maintaining the collected data (Rouse, 2019). This is done to ensure accessibility, reliability, and timeliness of the data. Also, it is important to protect the data. If the wrong people get access to it, it will harm your business and its reputation.

After the data has been collected and organized, data products are being used to generate value. Data products take data as input and provide useful insights as output. This is done by analyzing the data set and generating interesting insights. The output depends on the type of data product that is being used. Most data products are used for benchmarking, recommending, or forecasting. These insights are then used as a basis for the decision-making process. It is important that the people who work with the data product have enough experience in order to maximize its potential.

However, often the data is not the only input for the decision-making. As mentioned in the introduction, sources like experience and literature are valuable as well ("education is important", 2016). The level of experience someone has is based on a combination of his/her education and the number of years he/she has been active in this specific branch. Regarding literature, the level of validation is a key factor to measure its quality (Dellinger, 2005).

The decision-making itself consists of several variables like identifying alternatives and gathering relevant information (University of Massachusetts. (n.d.). This information can be used to determine the values of the choices (how important are the consequences?). Another important factor is the preference of the decision-maker since this will influence the decision to some extent. Sometimes there are multiple decision-makers, which makes it helpful to indicate the number of people involved in the decision (Mazurek, 2015).

Once the decision has been made, last step is to evaluate the decision outcomes. Did the decision that was made turn out to be the right one? It appears that the decision quality is based on all the previous dimensions. The quality of the decision outcome depends on the quality of the decision making, which in turn is a result of a combination of data products and experience/literature. However, the quality of the data products is determined by the quality of the data management, which in turn depends on the quality of the data. Even though it might take a while to see the results of the decision, this step can provide important feedback about the decision-making process.

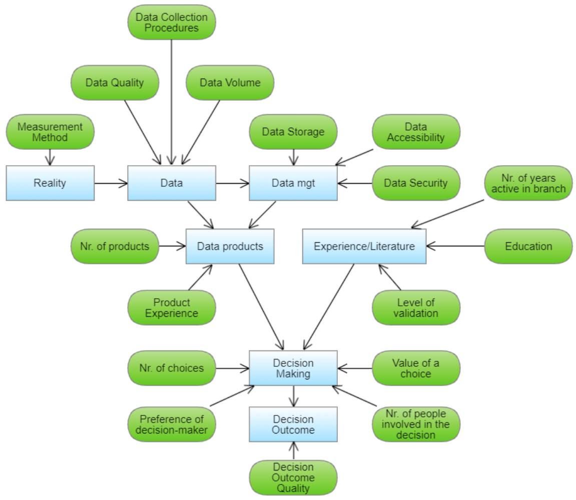

Figure 5: causal model for DDDM

4.5 Summary

When analyzing the literature regarding data-driven decision-making, several important components stick out. These serve as a basis for the decision-making process:

- Reality
- Data
- Data management
- Data products
- Decision-making
- Decision outcome

Now that the most important components in the data-driven decision-making process have been identified, lets include the variables that influence those components. To maintain a clear overview, two tables are provided that discuss for each variable how it can be measured. Table 2 shows the dimensions together with the variables by which they are influenced. For information on how to measure the variables, see table 5 in Appendix A.

DIMENSION	VARIABLE NAME
Reality	Measurement method
Data	Data quality
	Data collection procedures
	Data volume
Data management	Data storage
	Data accessibility
	Data security
Data products	Nr. of products
	Product experience
Experience/Literature	Nr. of years active in branch

	Education
	Level of validation
Decision-making	Nr. of choices
	Preference of decision maker
	Value of a choice
	Nr. of people involved in the decision
Decision outcome	Decision outcome quality

Table 2: variables per dimension

5. RESULTS

5.1 Preparation of the case study

Two dimensions of the causal model have not been converted into a level of the maturity model: 'Experience/literature' and 'Decision outcome'. The first was not included as a separate level since it does not relate to the data-driven aspect of the DDDM maturity model. However, it does influence the decision-making as most decisions do not fully rely on data but are based on experience as well. Therefore, the decision was made to include experience as a variable of the maturity model instead of being a dimension on its own. Furthermore, decision outcome is a way of confirming that all levels have been completed rather than being a level on its own. If all levels have been implemented correctly, the decision outcome should (almost) always be positive.

Another important difference is the change of the first level from 'reality' to 'awareness'. In theory, reality would be a good starting point for the process since every company can relate to it. Therefore, it makes sense to include it in the causal model. However, in practice, just reality is not enough to start the DDDM process. A specific mindset and more knowledge about the need for change is required. Conversations at Coulisse underlined the key role of this first level in the decision-making process by mentioning that awareness is required to trigger the process in the first place. Awareness indicates an understanding of the urgency for the need for change/improvement. Therefore, it was decided to not include 'reality' as the first level, but rather 'awareness'.

The variables related to the specific levels are also based on the causal model. However, after gaining more insights from the practical perspective, it was decided upon to leave out or replace some of the variables that are less relevant for businesses. This was done based on several meetings that took place with the information manager of Coulisse.

Furthermore, the maturity model has been reshaped to a more specific use instead of applying it to the whole company. At first, the goal was to measure the maturity level of a company regarding data-driven decision-making by means of a quantitative maturity model. However, problems were encountered on how to score the model since most variables are too process specific, making it hard to provide a score for the company in general. After discussing this issue with employees at Coulisse, it was decided that a more specific (qualitative) maturity model focusing on one single process is more valuable. Subsequently, several variables have been replaced by more specific ones.

The updated levels and variables are found in table 3. Since the first level has been changed from reality to awareness, the related variables obviously had to change as well. Regarding the data products level, two new variables have been added: product availability and type of products. The first has to do with the question whether all employees can access the data product and its results or just the person that takes the decision. Type of products refers to what kind of products are used: benchmarking, recommending, or forecasting. Last, the variables on the decision-making level have been adjusted as well since the maturity model is now applied to a specific process instead of the whole company. This change required the use of more specific variables. More details on the variables are given in appendix A.

Nr.	LEVEL	VARIABLE NAME
1	Awareness	Knowledge leading to awareness
		Need for change
2	Data	Data quality
		Data volume
		Data collection procedures
3	Data management	Data storage
		Data accessibility
		Data security
4	Data products	Product availability
		Product experience/ skills
		Type of products
5	Decision-making	Identify alternatives
		Gather relevant info
		Preference of decision maker

Table 3: Levels with their related variables

The order of the levels is important because of the way they influence each other (as can be seen in the causal model in section 4.4). Goal is to complete all levels starting with the lowest. A level is completed when all the related variables are implemented correctly. Only then you can move on to start working on the next level. To implement the variables, you should start at the Awareness level and end at the Decision-making level (see figure 6).

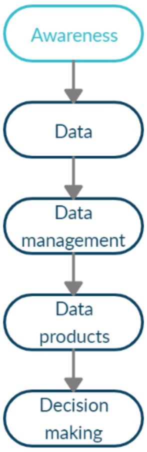

Figure 6: implementation part of the final maturity model

It is important to first think about what the levels and their variables should look like before immediately starting to implement the steps of the maturity model. This design process starts with awareness (see figure 7). You need to have understanding about the process as well as have knowledge about potential ways to improve the process. If this awareness is not in place, it makes no sense to start designing and implementing the next steps. However, if you know there are opportunities to improve the process, the model can be applied starting off with designing the decision-making level. You need to have a clear understanding of the goal: what decision needs to be made? Only once this is clear it is time to introduce the remaining levels. Start by thinking about what data products are required to solve the problem. Subsequently, think about how the data should be managed to efficiently use those data products. And last, consider what data is needed as input for the data products.

Now compare the results of your design with the current situation in your company. Identify what variables need to be changed and implement them in the right way.

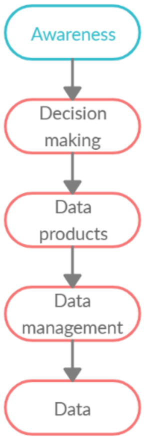

Figure 7: design part of the final maturity model

By combining the design and the implementation part of the maturity model, we end up with an interesting loop (see figure 8). This drastically changes the way in which to apply the model. Until now, you first had to identify all the variables that were not correct yet, then design all these variables in the correct way, and subsequently implement these variables according to the sequence of the levels of the model. The challenge would be that you had to design all the variables before implementing them. By doing so, it is hard to predict the consequences of designing a variable in a particular way. However, now that we have a loop, this process changes. First, you still must identify all the variables that are not correctly implemented yet. Next, you can start by designing the variables of all the levels and subsequently implement the variables of the first level. Now, the loop allows us to check if the design of the other levels needs to be changed because they might be affected by the consequences of implementing the first level. If they are, update the design of those levels and variables. If not, simply continue to implement the next level. This process can be repeated until all levels are implemented in the right way.

Once a process is successfully completed (= all the levels are correctly implemented), it is likely new insights have been discovered which lead to a new decision process. This way the maturity model can be repeatedly used by means of small increments. Aiming for large improvements will get you lost in the process while wasting a lot of time and money whereas aiming for smaller improvements helps you to maintain a clear overview and therefore to make more efficient progress. This process of improving over and over with small steps is called continuous improvement (Paipa-Galeano et al., 2020). The goal of continuous improvement is to increase efficiency and to reduce waste. Advantages are easy implementation and low costs compared to radical changes (Imai, 2012; Singh & Singh, 2015).

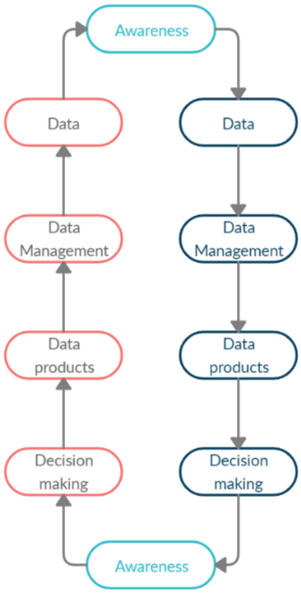

Figure 8: loop including both design and implementation

For every process that you apply this model to, the information optimization process (see figure 10) is relevant as well. In the case study, the model created in this section will be applied to several processes (see table 4). Some processes will be easy since the required data is available. This means that the first step from the information optimization process (descriptive analytics) can be performed relatively easy. For other processes, this data is missing. So, in order to perform descriptive analytics, first you need to create data that you can analyze (for example customer profiles). Next, before you can start forecasting, it is important to do diagnostic analytics first: analyze why certain type of customers order certain products. Only once you know the logic behind the data you can start forecasting.

5.2 Case Study

This part of the paper is about a case study which is performed to illustrate the intended use of the model that was constructed earlier in this paper. At the same time, it serves as a validation. As mentioned before, a specific process in a company is analyzed. In fact, to better validate the model, two processes were analyzed: one successfully completed process and one process that is not finished yet because it encountered challenges.

Both processes are about forecasting the number of products that need to be in stock. First process is about the relationship between existing products and existing customers. If the customer is known, and the products that can be ordered are known, is it possible to forecast how many products need to be in stock to anticipate on potential future orders and therefore to both keep the customer satisfied and keep the costs low? This process is successfully implemented by now. The other process deals with the question what happens when the factors of customer and products are unknown. This will make things more complicated.

The goal of the case study is to check whether the successfully implemented process followed the steps of the data-driven decision-making maturity model, and to identify on what level the other process encounters difficulties.

5.2.1 Introduction

The case study will be performed at Coulisse. This is a SME with about 180 employees. Coulisse is a global specialist in window coverings. The company was founded in 1992. Its employees are young and ambitious, resulting in a culture which is open to innovation and the use of new technology. However, the company is growing at a high pace over the last years, and now it starts to encounter challenges because of that growth.

The data collection during the case study consists of a combination of interviews and observations. Interviews have been conducted with four employees. They are all from different departments but worked together during the second project. Some of them have also worked on the first project, but most of them joined Coulisse when the first process was finished (almost all employees that worked at Coulisse during the first project are no longer active). The observations have been gathered during the time I was doing research at Coulisse. These observations include general knowledge about the company, analysis of data sources, but also discussions and conversations with employees as well as presentations used during internal group projects.

First interview is with an employee from inside sales who can be regarded as the initiator of the projects. Next interview is with the supply chain manager who has implemented the projects. Third interview is with a business controller who delivered the required data for the projects. Last interview is with the information manager. Because he was involved in the whole process, he was able to answer all questions that were left at this point.

The order of the interviews is not random but done with a purpose. When analyzing the different stakeholders of the projects by means of the model that has been created, it turns out that the employee from the first interview has knowledge about the decision-making and awareness levels since she initiated the process. Furthermore, the employee from the second interview has knowledge about the decision-making level and the data product level. The employee from the third interview is all into the data management and data, whereas the employee from the fourth interview can help with general questions about any level.

The interviews were semi-structured. Beforehand, key topics were written down that needed attention during the interview. However, no list of questions was put together as fully structured interviews tend to do.

5.2.2 DDDM maturity model in practice

The first interesting result that can be derived from the interviews and observations is that the two processes that were discussed can be split into four. Instead of only differentiating between known customers/products and unknown customers/products, we now get into more detail (see table 4). Making this distinction is important since it tells us that each process in table 4 is different and therefore needs its own approach. There is a difference in both the difficulty of a solution and in the frequency of the processes. Data regarding orders from customers at Coulisse showed that the combination of known customer / known product appeared more often than the combination of unknown customer / unknown product. However, even though the latter might not occur as often, its impact could be larger.

Recall that the goal is to be able to forecast how much stock is needed to provide all customers with the products they need, while at the same time keep the costs as low as possible. Even though dealing with a new customer is somewhat the same as dealing with a new product, there is a difference in how to solve these issues.

Customer/Product	**Known**	**Unknown**
Known	Easy	More difficult
Unknown	More difficult	Hard

Table 4: the processes used in the case study

Table 4 consists of four processes categorized into 'Easy', 'More difficult' and 'Hard'. A process is indicated 'Easy' when all required historical data is already available. 'More difficult' means that only part of the required historical data is available, and 'Hard' means that no historical data is available. In the latter cases, you need to make sure that you come up with some (additional) data set in order to perform descriptive-, diagnostic-, predictive- and prescriptive analytics (see figure 10). Lacking sufficient relevant data is what makes these processes so hard (also mentioned in section 3.2 as an important reason why it would be difficult to change to data-driven).

When looking at the overview in table 4, it makes sense to start with the easiest process. Once this is implemented, you can move on to the more difficult processes. At Coulisse, the process of forecasting the required level of stock when dealing with a known customer and a known product has been implemented successfully by now. This process is most easy since all the data is available. The customer has placed orders before so you can make use of their order history. The product has been in stock for some time which means that you can make use of the order history of that specific product. When combining these data sources, it is possible to forecast how many products you approximately need to have in stock in the near future. Since this project has been finished successfully, let us try to validate our model by applying it to the project (using the loop in figure 9 as guideline).

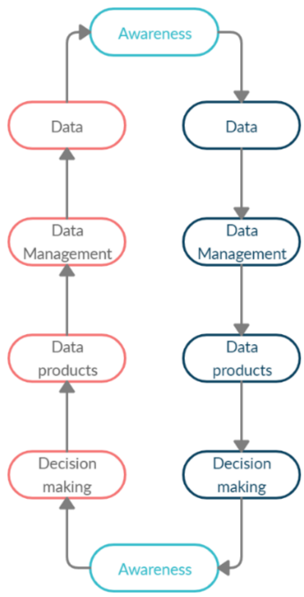

Figure 9: loop when applying the DDDM maturity model

Recall that applying the model consists of three steps: design the levels of the model how they should be, identify what variables are currently not implemented correctly, and then implement them in the right way.

Applying the model to the process of known customers and known products encounters some challenges due to the fact that this project took place about six years ago. Most employees that worked on the process are no longer at Coulisse, and during that time I was not there to observe. However, sufficient information was gathered to be able to apply the model to the process.

First of all, the interviews showed that the awareness was in place since there was a clear need for change. Until then, this process was mostly done by hand with the help of Microsoft Excel. However, more often limitations of Excel were causing challenges. Also, until this point in time, the number of products and stock locations have been relatively small. But now these numbers started to become so large (data showed me somewhere between 5000 and 7000 products, of which almost every single product has its own unique product location) that everyone agreed it was time to automate the process. The only question was how. About six years ago, this question was the motivation for starting a project with the goal of being able to forecast the required level of stock when both customer and product are known.

This already tells us that the design of the decision-making level was in place since the goal of the project was clear: to forecast the required number of stock when dealing with a known customer and known products; Furthermore, alternative solutions had been identified (like simply increasing the overall level of stock) and relevant information was gathered. Because the goal was to do forecasting, a data product would be needed that could do the job. After exploring some of the alternatives, they soon came up with a product called slim4. This product requires data from historic orders of a specific customer, which it then uses to forecast the next order. By doing so, it decreases the number of times an 'out of stock' and 'over stock' happens. 'Out of stock' meaning that a customer orders more products than expected (causing Coulisse to have insufficient stock), and 'over stock' meaning a customer orders less products than expected (leaving coulisse with too much stock). If this would be the product they were going with, data regarding the previous orders of a customer would be needed. Since this data is not hard to get, this would not cause major problems or require huge changes. Only currently with the covid19 situation it is difficult to predict how all this will impact the order amount of customers, since never before they have experienced a situation like this which means there is no historic data available to compare to.

The paragraph above describes the results of the design part when applying the model to the process of forecasting the level of stock when dealing with known customers and known products. Implementing the proposed solution would require changing variables on several levels. The first level (awareness) was correct, but the second level (data) already would require changes since new data was needed for the data product (namely customer history and product history). This means that at this point in time the process would be identified as being on level 1, since the second level was not implemented correctly yet. After fixing the variables of the data level, the process would be identified as level 2 since the data management level needed to be changed as well. The data needed to be stored in such a way that the right information would be accessible for the right people and data products. Once this had been fixed, the data product (in this case Slim4) had to be implemented in order to complete level 4. Also, the product needed to be available to employees that had the right skills. Last step was to use the results from Slim4 to start making real decisions about the level of stock. Data shows that, after a period of getting used to working with Slim4, the number of 'out of stock' and 'over stock' indeed declined. Because at the end all levels were successfully implemented in the right way, and the initial goal to forecast the level of stock was achieved, we can say that this process can now be identified as level 5 which is the maximum level in the model.

After this project was successfully finished, new challenges were encountered (this is exactly what was meant when mentioning continuous improvement). What if a new product is released, and some potential new customer might be interested in this product? How can you approximate the number of products you need to preventive have in stock? This brings us to the second process which is looking for the required level of stock when dealing with a new product and a new customer. According to the interviews, the awareness was in place since time after time the employees were encountering customers complaining about products that were not in stock. Often the reason was that a new customer would come up with an order that was larger than expected so that either this new customer was not able to get all the products, or this new customer took all the products that were produced for other customers. Or sometimes, the new customer would place an order that was smaller than expected, leaving Coulisse with too many products in stock (which costs money). Observing the data showed me that indeed still 'out of stock' happened, although drastically declined after the introduction of Slim4 during the previous project.

These issues raised a certain level of awareness that led to the start of a new project. The goal of this project was clear: forecast the required level of stock when dealing with a new customer and a new product. When applying the model, first task is to design the levels. What is needed in order to achieve the goal? After generating awareness, the design phase starts at the decision-making level, by identifying alternatives and gathering relevant information. However, when asking the employees during the interviews what potential solutions have been thought of, they could often think of maximum one or two possible solutions, which often differed from the other employees. This indicates that the alternatives have not clearly been identified and discussed yet. Currently, several different solutions have been proposed during the interviews:

- Discuss with the customer how many products they think they are going to need. Problem is that often the customer does not know either and that they do not want to be obligated to order the number of products that they estimated they will need.
- Wait until the customer has placed three orders. Three orders are the minimum input that slim4 currently needs to forecast the next order. Problem is that these first three orders need to be forecasted as well.
- Come up with customer profiles and compare the new customer to the behavior of other customers within the same profile. The same can be done with products.
- Maybe the solution is not data related. What if forecasting is not the solution. Maybe it is sufficient to increase the average stock level. Or maybe it is not too big of a deal to have to tell the customer it is going to take a little longer. If the consequences are not terrible, this might save a lot of costs.

Since many current potential solutions have disadvantages attached to them, and no specific one has been identified as most suitable yet, this step requires more attention. The goal of the project is clear, it is just not known yet how to reach it.

Regarding the data product level, they use slim4. Slim4 is currently not suitable to do the job since it requires three orders of the same customer as input before being able to forecast the next order. Therefore, designing the data product level results in the choice to either make sure to somehow gather the required data that slim4 needs (since the skills and experience are already in place for using slim4), or to investigate potential other data products. But this choice depends on what alternative is selected at the previous (decision-making) level.

According to both interviews and observations, the data management level needs some work as well. Compared to the other project, which only made use of the order history of customers, now many more data sources would need to be used and combined. During the interviews it became clear that the creation of a data warehouse is in progress, which will be an important step in the right direction. After looking into it, I found that Coulisse currently makes use of about 15-20 applications that each have their own database. They are indeed working on a data warehouse to gather all the data. Right now, they focus on the transactional data (like purchase orders and sales orders). Once that is done, other kinds of data will follow (like customer data, product data).

The data level probably causes the most trouble currently, simply because the required data for Slim4 is not available. Therefore, research is needed to identify what data is needed (for example customer/product profiles), and how to gather this data.

Once this design phase is completed, next step would be to identify what variables need to be changed and then to implement the design. In this scenario however, the design is not clear yet. First, a potential solution needs to be identified as most suitable, so that the other levels can be designed in line with this solution. Therefore, the design needs more attention because it would be impossible to identify what variables need to change if you do not know yet what to change them to. Only once the full design is clear, it is time to proceed to identifying what variables need to change.

If we had to identify the current level of this second process, it would be level 1. Awareness is in place, so the first level is completed. However, most likely new data needs to be gathered (we are not sure because this depends on what alternative is selected as best solution). This means that the second level would not be implemented successful, which keeps the project at level 1. Therefore, to increase this level, it is recommended to first complete the design phase.

5.3 Analysis

Even though the first process was implemented successfully, it currently is not performing well due to the covid19 situation. Because of covid19, customers drastically change their orders. This is messing up the forecast process, since no historic data is available of how orders are impacted during a pandemic. This is interesting since this issue has many similarities with the second process: a new customer means no historic data of that customer, and a new product means no historic data of that product. So once a solution to the second process is available, this might help improve the first process.

Figure 10 shows the process that Coulisse is working on. Their current goal is to get to the level of predictive analytics: 'what will happen?'. However, as can be concluded from the interviews and observations, they so far only successfully implemented the descriptive analytics: 'what happened?'. Although they are working on the diagnostic analytics (why did it happen?), they are already trying to achieve level three. This indicates one of the most important reasons why the second process (unknown customer/unknown product) is encountering so many issues: the steps that Coulisse is trying to take at once are way too large. Applying the model to the second process revealed that many parts of the model were not implemented correctly yet. Therefore, it would be best to first focus on these issues, starting with the lowest level.

According to the analysis above, problems were encountered on the levels of decision-making, data product, data management and data. But when using the model, first we should look at the design part. So, when looking at the sequence of the levels in the design part, it is best to start working on the decision-making level since the awareness level is already correct. The decision-making level consists of several variables that should all be clearly answered. Once the decision-making level is successfully designed, it is time to move to the data product level and do the same. This should be repeated until all levels are correctly designed. Because only then it is possible to identify what variables need to change and to implement them in the correct way. Important is to keep the steps small. Iterate over the model many times whilst making small changes. This way it is much easier to maintain a clear overview and to know what step is next compared to immediately trying to achieve something that is too challenging.

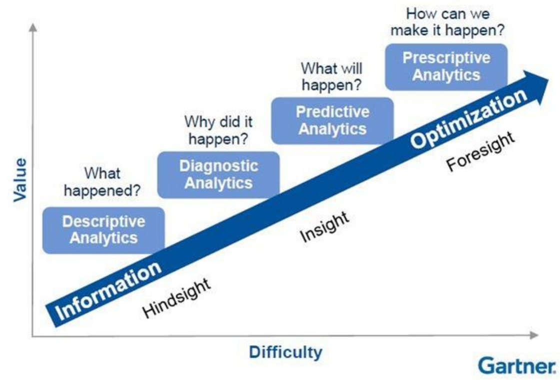

Figure 10: information optimization (source: Gartner.com)

So, to conclude this case study, the step to forecasting unknown customers/unknown products would be too large for Coulisse. It is still unclear what the best solution would be. This might be solved by putting in lots of time, but it would be best to apply the model in small incremental steps. The model indicates that currently issues occur on multiple levels. So instead of immediately deciding to do forecasting for the combination of new customers and new products, it is better to first slow down and look at just new products or just new customers. So instead of immediately rushing to the bottom right in table 4, first the bottom left and top right need to be solved. And most likely the steps need to be even smaller to get the best quality progress, like first analyzing the behavior of current customers and products. This would be a great moment to start making optimal use of continuous improvement by iterating the process over the model repeatedly, whilst in the meantime making small adjustments. This solves the current problem of not knowing what the next step is, while at the same time it optimizes the process by solving the issues piece by piece.

5.4 How to implement DDDM

This section will provide some clear, practical guidelines on designing the project when starting to apply DDDM to a business. First step is to raise awareness and create a data-driven mindset. This needs to happen before the project can really start. You need to make sure that people see the value of data-driven since that opens up opportunities to start new projects. Coulisse is already changing their culture, which makes it possible for them to introduce more (challenging) data-driven related projects.

Once this awareness (the first level of the model) is in place, you need to decide on the end goal of the project. Make sure this goal is clear to everyone. Now you can start on designing a potential implementation. Use the model to do so and stick to it. Since the awareness level is in place by now, use the decision-making level as starting point to identify alternative ways to reach the end goal, and gather relevant information regarding those alternatives; once completed, move on to the next parts of the model. Use the model in small iterations to enable continuous improvement: design and implement small changes and look at its consequences; then change your overall design if needed and start designing and implementing the next step. If you get stuck during the project, it is very likely that you know what needs to be changed (just look at the variables of the model and decide whether they are correctly in place yet).

When applying this model, keep in mind that it is important to have open discussions between the involved departments. This is an important lesson that was learned when using the model during the case study. Make sure to clearly agree upon the responsibilities of each department to avoid irritations.

6. DISCUSSION

Many types of maturity models exist in literature, of which several are related to data. These existing data-related maturity models are used when dealing with specific parts of the DDDM process. Data quality maturity models for example ensure the quality of the data, which is one of the variables of the DDDM maturity model constructed in this paper. Big data maturity models help to structure the processes regarding big data and to determine where to start. Data management maturity models help manage the data, which is also an important part of the DDDM process. Point is that in the end, they all help to reach the same goal: DDDM. However, no maturity model currently exists that guides you in the overall DDDM process itself. Also, according to businesses, it is difficult to determine where to start and it is easy to lose track during the process because of all the existing models. This paper adds to theory by constructing a simple, yet effective DDDM maturity model to guide businesses in their change to a more data-driven decision-making process.

According to Gregor & Hevner (2013), design science research can be divided into four categories, namely Invention (new solution for new problem), Improvement (new solution for known problem), Exaption (known solution for new problem) and Routine design (known solution for known problem). This paper would be part of the Exaption category since maturity models already do exist in many disciplines. However, this paper extends the use of maturity models into DDDM. Current literature provides many kinds of maturity models, but surprisingly no DDDM maturity model could be found. This indicates a gap in the available theory which is filled by the research described in this paper.

It is best to use this DDDM maturity model whilst keeping the goal of continuous improvement in mind. The idea of continuous improvement is to improve a process in small steps instead of aiming for one large step. Main advantages are an increase in efficiency and a reduction of waste. Some authors already made the connection between data-driven and continuous improvement. However, this paper adds to that by doing so specifically for DDDM. This paper shows how continuous improvement can be unlocked and how it should be used when applying the DDDM maturity model. This can be achieved because of the loop that has been created in section 5.1. When applying the model to a process, it is best to iterate repeatedly over the levels of the model whilst making small improvements. For every new idea for improvement, go through the design part of the loop to create an idea of what the levels should look like. Use the model to identify what levels cause issues. Once this has correctly been done, move on to the other part of the model which consists of implementing the design. Start by fixing the first level that encounters issues; next, go over the design part and see if the consequences of the implementation are still in line with your design. If not, edit the design or reimplement if needed; if it is still in line, continue with implementing the next level. Once all the issues have been solved, and the whole loop has correctly been finished, it is likely that a new idea for improvement has come to mind. Now the model can be applied again.

As mentioned in the introduction, to identify whether the model would be valid and therefore meaningful to practice, it is important to see if different processes get stuck on different levels of the model. If every process would run into issues on the same level, this would indicate some weakness of the model. The results of the case study showed that both processes encountered issues on multiple levels, which indicates that the model is working according to its purpose. Also, regarding the successfully finished process, at first several variables were identified that required a different implementation. Only once all these variables were correct, the process could be regarded as successful. This tells us that the variables that are included in the model seem to be relevant.

However, it should be noticed that during the case study the model has been applied to only two processes. To further validate the model, it needs to be applied to more processes.

The first process (known customer/known product) has at the end been identified as maximum level, indicating that it is possible to successfully implement all levels. However, it is unlikely that this happens at the first try; almost always at least one iteration over the model is required (further research would be needed to substantiate this claim). The results regarding the second process (unknown customer/unknown product) indicate that if a project gets stuck at some point, the model can be used to identify the problems in a simple but effective way. First determine what level needs to be fixed, then what variable(s). Because of the use of continuous improvement, these challenges can be divided into smaller steps that are easier to implement. By subsequently checking if the consequences are still in line with your overall design, you can quickly anticipate if needed. Doing so prevents the unnecessary investment of lots of time, money, and other resources if things go wrong.

Applying the model in a case study is important to validate the model in practice. Especially since the model was largely constructed based on literature and theory. However, it is important to take the context of validation into consideration. During the case study, the model was used at Coulisse, a SME which is currently focussing on raising awareness for using data. Until now, only some employees were familiar with its potential value. Majority of the employees, including the board, would not recognize the importance of data. However, this is changing now. The culture at Coulisse is open to innovation, which helps to stimulate a data-driven mindset. Having such a culture and mindset in your company is crucial for using this model, since the important first level and starting point of the measurement instrument is about awareness. You need to acknowledge a need for change, and therefore you must know the potential of using data. According to my observations, many businesses hear hype words like forecasting and AI, and immediately want to apply this to their own company. However, the important lesson this DDDM maturity model teaches is to use small steps when growing. Otherwise people will get lost in the process, which results in a lot of extra costs.

Because of this, the model will be most useful for companies that are transitioning their culture into a data driven one. This simple yet effective model is practical and helps you as a business to determine where to start in the DDDM process. This is an important distinction between this paper and current literature regarding DDDM. Even though DDDM is no new topic, and several guidelines are available on how to change to data-driven, businesses indicate that it is hard to use them because of not knowing where to start. This paper however acknowledges this problem by focussing on the first step of this process (the awareness level).

As mentioned, the model has been validated during a case study by applying it to two processes regarding the level of stock of a business. These processes have not been chosen at random, but rather because of their generalizability. Many businesses have this problem of new customers/new products meaning they can learn from the outcomes of the case study. Furthermore, this challenge of new customers/new products can be generalized to the challenge of demand and offer since a new customer can be regarded as demand and having a new product in stock can be regarded as offer. Therefore, even businesses that do not have to deal with new customers or new products can still relate to this research.

Most likely, the model is useful for other processes as well, as long as the end of the process results in data-driven decision-making. When all levels are implemented correctly, you must be able to use them as input for your decisions. The processes in the case study were for example about forecasting the level of stock. Once all the levels were successfully implemented, the employees at the supply chain department were able to use these insights for their ongoing decisions of how much stock would be required.

Let us now come back to the hypothesis stated in the introduction about the DDDM maturity model being useful for every company. Now that we have seen the results of the case study, it seems that the model is useful for all kinds of businesses. So, in line with the hypothesis, it does not depend on the type of business or the sector you are in. However, it should be noted that you need a culture and mindset that is open to data driven. Therefore, it is not proven yet that the model is useful in every context.

Another possible limitation of the model could be regarding its variables. When applying the model during the case study, it seemed that the included variables are relevant. However, there might be more variables that influence the DDDM process which are currently not included. Furthermore, it is not proven that the model is applicable to every context. As mentioned earlier in this section, the model was validated in a context that involves a certain level of awareness. To examine in more detail whether the model can be applied in other contexts, more in-depth research is needed.

Almost all of the components that are included in the DDDM maturity model are based on literature and confirmed in practice. Only few of them are heavily based on practice, namely awareness and product availability. Awareness has been mentioned multiple times by now as being an important factor to initialize the process, even though it is not mentioned a lot in current literature. Furthermore, product availability is also not mentioned a lot in papers, but after talking to businesses it was decided to include it as a variable on the 'data product' level. This variable refers to the ability to make use of the product. Often, data products are licensed. And because companies want to save money, not every employee will receive a license even though it helps them if they do. Variables that are based on literature but were not proven relevant have been left out of the final DDDM model. These can be found in the causal model (see figure 5).

Next, although not proven, the model is likely to stimulate organizational learning. This concept means that a business gains experience, from which it can create knowledge. When a business improves a process by iterating over the model, it builds experience on how to increase efficiency and solve certain issues. This experience can then be converted into knowledge which is useful when implementing new improvements in the future. Especially the combination with continuous improvement creates many iterations over the model which result in many new experiences and learnings.

Furthermore, most of the literature will tell you about what the best way is to do things. However, doing empirical research shows you the challenges that come with the implementation. The case study showed us why it can be difficult to make use of DDDM, and why this aspect deserves more focus in literature.

First, it is clear that conditions exist that need to be fulfilled in order to make good use of the DDDM maturity model. Something that is somewhat hidden in literature. These conditions mainly consist of having awareness and a correct mindset in place. They are needed to trigger the process: you need to know about the potential of data-driven in order to use it to improve your processes. Furthermore, the added value needs to be clear. According to my observations and conversations with employees, board members do not want to switch to data-driven unless they are 100% sure that this will result in more value than the current way of working.

Next, as mentioned before, companies struggle to find the first step when switching to data driven. So, although many models exist in literature about the topic of data-driven, companies need models that focus on practice. It is therefore important to provide clear guidelines on how to start and improve in the DDDM process, something that should require more focus in current literature.

Last, the interviews performed during the case study show us something interesting about the challenge of implementing according to literature. When improving a process regarding DDDM, the difficulty is that multiple departments need to work together. This results in discussions between employees about what department has what responsibility. Current literature does not provide a lot of information about this specific issue, even though it is important according to the results of this research.

The key takeaways from the discussion are that this paper managed to construct a DDDM maturity model whereas such a model did not exist in literature yet. After applying the model in a case study, the results show some encouraging first signs of the model being valid. Furthermore, because the model was created in such a way that it consists of both a design and implementation part, it stimulates the appliance of continuous improvement by iterating over the model in small steps.

7. CONCLUSIONS

MRQ – What is a valid maturity model to measure and improve the level of data-driven decision-making?

See the model that was created in section 5.1 (see table 3 and figure 8). This simple, yet effective model is suitable for measuring the data-driven decision-making maturity level. Also, it can be used as a guideline to improve the current level. The case study shows positive signs of the model being valid, but further research is required to prove full validity.

SQ1 – How can a company become more mature regarding data-driven decision-making?
For specific processes by making sure the levels and variables of the data-driven decision-making maturity model (see table 3) are correctly implemented; In general for the company by making use of continuous improvement: making sure to not aim for goals that are currently too challenging, but rather make progress by means of small steps.

8. Limitations & Future work

As mentioned in the discussion, an important limitation to the DDDM maturity model is its level of validity. So far, the model has only been applied to two processes during a case study. However, to fully prove if the model is valid, it needs to be applied more often. Only then it can be assured that these are all relevant variables that influence the DDDM process. Further validation will also provide more insight into what context is required to make optimal use of the model.

Furthermore, this paper focusses on the data-driven part of the decision-making process. However, it does mention that experience and human judgement can still be important as well. This means that not every decision should be based on (only) data. Additional research would be helpful to gain more detailed insights in the role of experience and human judgement in the decision-making process. Will data-driven be able to substitute them in the future? If not, what kind of decisions cannot be substituted?

Last, it would be helpful to gain more in-depth insight in the value that DDDM brings. This paper does mention the impact data-driven has on the decision-making process as well as several advantages (based on literature). However, it would be more attractive for businesses to adopt data-driven if its value can be made specific.

REFERENCES

Althuizen, N., Reichel, A., & Wierenga, B. (2012). Help that is not recognized: Harmful neglect of decision support systems. *Decision Support Systems, 54*(1), 719–728. doi: 10.1016/j.dss.2012.08.016

Asadi Someh, Ida and Shanks & Graeme, "How Business Analytics Systems Provide Benefits and Contribute to Firm Performance?" (2015). *ECIS 2015 Completed Research Papers*. Paper 12. doi:10.18151/7217270

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering, 1*(3), 213-222. doi:10.1007/s12599-009-0044-5

Benbasat, I., & Zmud, R. W. (1999). Empirical Research in Information Systems: The Practice of Relevance. *MIS Quarterly, 23*(1), 3. doi:10.2307/249403

Berg, V., Birkeland, J., Pappas, I. O., & Jaccheri, L. (2018). The Role of Data Analytics in Startup Companies: Exploring Challenges and Barriers. *Lecture Notes in Computer Science Challenges and Opportunities in the Digital Era,* 205-216. doi:10.1007/978-3-030-02131-3_19

Brocke, J. V., Mathiassen, L., & Rosemann, M. (2014). Business Process Management. *Business & Information Systems Engineering, 6*(4), 189-189. doi:10.1007/s12599-014-0330-8

Brownlee, J. (2007). *Satisficing, Optimization, and Adaptive Systems* (pp. 1-5, Tech. No. 070305A). Melbourne, Australia: Complex Intelligent Systems Laboratory, Centre for Information Technology Research, Faculty of Information Communication Technology, Swinburne University of Technology. Retrieved June 11, 2020.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? *SSRN Electronic Journal*. doi: 10.2139/ssrn.1819486

CIO wiki. (n.d.). Capability Maturity Model (CMM). Retrieved May 25, 2020, from https://cio-wiki.org/wiki/Capability_Maturity_Model_(CMM)

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research, 70*, 379-390. doi:10.1016/j.jbusres.2016.08.011

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Boston, MA: Harvard Business School press.

Dean, J. W., & Sharfman, M. P. (1996). Does Decision Process Matter? A Study of Strategic Decision-making Effectiveness. *Academy of Management Journal, 39*(2), 368–392. doi: 10.5465/256784

de Bruin, T., Freeze, R., Kulkarni, U., & Rosemann, M. (2005). Understanding the Main Phases of Developing a Maturity Assessment Model. *ACIS 2005 Proceedings*.

Dellinger, A. B. (2005). Validity and the Review of Literature. *Research in the Schools, 12*(2), 41-54.

Demchenko, Y., Laat, C. D., & Membrey, P. (n.d.). *Defining architecture components of the Big Data Ecosystem*. Amsterdam: University of Amsterdam.

Divan, M. J. (2018). Data-driven decision making. *2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS), 2018-January*, 50–56.

Education Is Important, But Nothing Beats Experience In A Software Engineer. (2016, August 25). Retrieved June 04, 2020, from https://www.forbes.com/sites/quora/2016/08/25/education-is-important-but-nothing-beats-experience-in-a-software-engineer/

Edwards, W. (1954). THE THEORY OF DECISION MAKING. *PSYCHOLOGICAL BULLETIN, 51*(4), 380–417.

Eisenhardt, K. M., & Zbaracki, M. J. (1992). Strategic decision making. *Strategic Management Journal, 13*(S2), 17–37. doi: 10.1002/smj.4250130904

Ejbye-Ernst, D., & Jorring, N. (2017). Doing It Collaboratively! Addressing the Dilemmas of Designing Quantitative Effect Studies on Narrative Family Therapy in a Local Clinical Context. *Journal of Systemic Therapies, 36*(1), 48-66. doi:10.1521/jsyt.2017.36.1.48

Gottschalk, P. (2009). Maturity Levels for Police Oversight Agenciespetter Gottschalk. *The Police Journal: Theory, Practice and Principles, 82*(4), 315-330. doi:10.1350/pojo.2009.82.4.472

Gray, K. D. (n.d.). 5 Steps to Good Decision Making. Retrieved June 11, 2020, from https://www.corporatewellnessmagazine.com/article/5-steps-to-good-decision-making

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly, 37*(2), 337-355. doi:10.25300/misq/2013/37.2.01

Gregory, M., & Roberts, C. (2020). Maturing an Information Technology Privacy Program. *Evaluating Challenges and Opportunities for Healthcare Reform Advances in Healthcare Information Systems and Administration,* 125-143. doi:10.4018/978-1-7998-2949-2.ch006

Greiner, L. (1998, August 01). Evolution and Revolution as Organizations Grow. Retrieved May 26, 2020, from https://hbr.org/1998/05/evolution-and-revolution-as-organizations-grow

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19-30. doi:10.1037/1040-3590.12.1.19

Hearn, S. (2020, March 16). What do we mean by 'impact'? Retrieved May 27, 2020, from https://www.researchtoaction.org/2016/02/what-do-we-mean-by-impact/

Herbert A. Simon, A Behavioral Model of Rational Choice. *Quarterly Journal of Economics, vol. 69*, pp. 99-118, 1955

Herbert A. Simon, Rational choice and the structure of the environment. *Psychological Review, vol. 63*, pp. 129-138, 1956

Iivari, J. (2003). The IS Core - VII: Towards Information Systems as a Science of Meta-Artifacts. *Communications of the Association for Information Systems, 12*. doi:10.17705/1cais.01237

Imai, M. (2012). *Gemba Kaizen: A Commonsense Approach to a Continuous Improvement Strategy; Second Edition*. McGraw-Hill.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697-720. doi:10.1037/0003-066x.58.9.697

Kazanjian, R. K., & Drazin, R. (1989). An Empirical Test of a Stage of Growth Progression Model. *Management Science, 35*(12), 1489-1503. doi:10.1287/mnsc.35.12.1489

Khan, N. (2019, December 23). How to make SMEs more data-driven. Retrieved June 18, 2020, from https://digileaders.com/how-to-make-smes-more-data-driven/

Kleindienst, D. (2017). The data quality improvement plan: deciding on choice and sequence of data quality improvements. *Electronic Markets 27*:4, 387-398.

Klein, G. (1995). *Decision making in action: models and methods*. Norwood, NJ: Ablex Publishing Corporation.

Kuznets, S. (1965). *Economic growth and structure: Selected essays*. London: Heinemann.

Lahrmann, G., Marx, F., Winter, R., & Wortmann, F. (2011). Business Intelligence Maturity: Development and Evaluation of a Theoretical Model. *2011 44th Hawaii International Conference on System Sciences*. doi:10.1109/hicss.2011.90

Manyika, J., Chui, M., Brown, B., et al. (2011) *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute.

Maslow, A. H. (1954). *Motivation and personality*. New York: Harper & Row.

Mazurek, J. (2015). *On a Preference Analysis in a Group Decision Making* (pp. 1-11, Publication). Opava, Czech Republic: Silesian University.

Mcelheran, K., & Brynjolfsson, E. (2017). The rise of data-driven decision-making is real but uneven. *IEEE Engineering Management Review, 45*(4), 103-105. doi:10.1109/emr.2017.8233302

Merriam-Webster. (n.d.-a). Maturity. Retrieved May 14, 2020, from https://www.merriam-webster.com/dictionary/maturity

Merriam-Webster. (n.d.-b). Valid. Retrieved May 14, 2020, from https://www.merriam-webster.com/dictionary/valid

Mettler, T., and Rohner, P., "Situational Maturity Models as Instrumental Artifacts for Organizational Design", *ACM, 2009*, pp. 1-9.

Mintzberg, H., Raisinghani, D., & Theoret, A. (1976). The Structure of "Unstructured" Decision Processes. *Administrative Science Quarterly, 21*(2), 246. doi: 10.2307/2392045

Moreno, H. (2017, June 05). The Importance Of Data Quality -- Good, Bad Or Ugly. Retrieved June 22, 2020, from https://www.forbes.com/sites/forbesinsights/2017/06/05/the-importance-of-data-quality-good-bad-or-ugly/

Müller, O., Fay, M., & Brocke, J. V. (2018). The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics. *Journal of Management Information Systems, 35*(2), 488-509. doi:10.1080/07421222.2018.1451955

Navarro, F. J. (2014, October 11). Francisco J Navarro. Retrieved May 19, 2020, from https://customerthink.com/what-can-a-maturity-model-do-for-you/

Nolan, R. L. (1973). Managing the computer resource. *Communications of the ACM, 16*(7), 399-405. doi:10.1145/362280.362284

Paipa-Galeano, L., Bernal-Torres, C.A., Agudelo-Otalora, L.M., Jarrah-Nezhad, Y., & González-Blanco, H.A. (2020). Key Lessons to Sustain Continuous Improvement: A Case Study of Four Companies. *Journal of Industrial Engineering and Management, 13*(1), 195-211. https://doi.org/10.3926/jiem.2973

Paulk, M.C., "A History of the Capability Maturity Model for Software", *ASQ Software Quality Professional, 12*(1), 2009, pp. 5-19.

Paulk, M.C., Curtis, B., Chrissis, M.B., Averill, E.L., Bamberger, J., Kasse, T.C., Konrad, M.D., Perdue, J.R., Weber, C.V., & Withey, J.V., "*Capability Maturity Model for Software*", Carnegie Mellon University, Pittsburgh, PA, 1991

Pöppelbuß, J. & Röglinger, M. (2011). WHAT MAKES A USEFUL MATURITY MODEL? A FRAMEWORK OF GENERAL DESIGN PRINCIPLES FOR MATURITY MODELS AND ITS DEMONSTRATION IN BUSINESS PROCESS MANAGEMENT. *ECIS 2011 Proceedings*. Paper 28

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and DataDriven Decision Making. *Big Data, 1*(1), 51–59. doi: 10.1089/big.2013.1508

Pulkkinen, M. (2013). An Integrated Model for Evaluating ICT Impact in the Education Domain. In *European Conference on IT Management and Evaluation (ECIME).* Gdańsk, Poland.

Rosemann, & Vessey. (2008). Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks. *MIS Quarterly, 32*(1), 1. doi:10.2307/25148826

Rouse, M. (2019, October 30). What Is Data Management and Why Is It Important? Retrieved May 27, 2020, from https://searchdatamanagement.techtarget.com/definition/data-management

Sein, Henfridsson, Purao, Rossi, & Lindgren. (2011). Action Design Research. *MIS Quarterly, 35*(1), 37. doi:10.2307/23043488

Selladurai, R., Hobson, C., Selladurai, R. I., & Greer, A. (2020). *Evaluating challenges and opportunities for healthcare reform*. Hershey, Pennsylvania (701 E. Chocolate Avenue, Hershey, Pennsylvania, 17033, USA): IGI Global.

Singh, J., & Singh, H. (2015). Continuous improvement philosophy – literature review and directions. *Benchmarking: An International Journal, 22*(1), 75-119. doi:10.1108/bij-06-2012-0038

Smits, D., & van Hillegersberg, J. (2015). IT Governance Maturity: Developing a Maturity Model Using the Delphi Method. *In 48th Hawaii International Conference on System Sciences (HICSS)* (pp. -). Kauai, HI: IEEE. https://doi.org/10.1109/HICSS.2015.541

Streifer, P. A., & Goens, G. A. (2004). *Tools and techniques for effective data-driven decision making*. Lanham, MD: ScarecrowEducation002E

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39-60. doi:10.1016/0167-2681(80)90051-7

Thiess, T., & Müller, O. (2018). Towards design principles for data-driven decision making - An action design research project in the maritime industry. *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change*, *ECIS 2018*.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131. doi:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323. doi:10.1007/bf00122574

University of Massachusetts. (n.d.). Decision-making process. Retrieved June 05, 2020, from https://www.umassd.edu/fycm/decision-making/process/

Vohra, G. (2016, August 16). Here Are The Benefits of Data-Driven Decision Making. Retrieved June 18, 2020, from https://www.entrepreneur.com/article/280923

Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and Software*

Technology, 54(12), 1317–1339. doi: 10.1016/j.infsof.2012.07.007

Wu, L., Hitt, L. M., & Lou, B. (2016). Data Analytics Skills, Innovation and Firm Productivity. *SSRN Electronic Journal*. doi:10.2139/ssrn.2744789

APPENDIX

A – variables explained

The table below explains the variables of both the causal model of section 4.4 (see also table 2) and the maturity model of section 5.1 (see table 3) in more detail. The inclusion of these variables has all been based on literature. References to the literature can be found throughout the paper, especially in section 4.4 and 5.1.

Variable	How to measure
Measurement Method	Identify what data is being collected. This is done by consulting both the person that gave the order to collect the data and the person that collects the data.
Knowledge leading to awareness	Knowledge is required about potential improvements to current processes. Can be determined by analyzing the way of thinking of employees.
Need for change	Employees need to recognize that things need to change before triggering the model. Can be determined by talking to employees.
Data Quality	Determine the quality of the data. Can be achieved by checking if the collected data fits its intended use. For example: is the data correct and complete?
Data Volume	This is the amount of data that has been collected. Can be found when consulting the database(s). There should not be a data overload, but too little data is less suitable for using data products. Determine how much data is needed for the project.
Data Collection Procedures	This refers to how the data has been collected. Have reliable and valid procedures been used? Can be identified by studying the measure procedures as well as consulting the person(s) that use these procedures.
Data Storage	Where is the data being stored? Ideally, all data is gathered in one place (data lake / data warehouse). If not, the different data locations should be efficiently connected. Check how the data is stored.
Data Accessibility	The data should be easily accessible for the employees that need it. This regards internal data as well as external data. The latter might cause problems though. Consult employees that need access.
Data Security	The level of security of the collected data. Not just anyone in the company should be able to access the data. And definitely not anyone outside the company. Can be measured by checking database settings and by consulting security experts.
Nr. of products	Identify how many data products are in place. These can for example include data mining or data analytics.
Product availability	Can all employees that need to access (results of) the data product? Can be measured by checking the rights for every employee.
Product Experience/ skills	Determine the level of experience that employees have with the data products. Have they been using them for years or were these processes just implemented?
Type of product	What is the purpose of the data product? Check if benchmarking, recommending or forecasting.
Nr. of years active in branch	This regards the experience of the decision-maker. The number of years that this person has been active in this specific branch says a lot about his/her experience. Can be determined by consulting this specific person.

Education	Identify the level of education the decision-maker has had during his life. The more/higher education, the more knowledge this person has which results most likely in a more reliable decision.
Level of validation	The level of validation concerns the literature. When using certain research, one should consider its validity. Can be determined by studying the research, reading reviews and checking how many papers are referencing to this study.
Nr. of choices / identify alternatives	When making a decision, all possible options should be considered in order to not miss out on the best choice. Brainstorm about all potential alternatives.
Preference of decision-maker	Almost always, a decision-maker is affected by his/her own opinion/preference when taking a decision. This might be a disadvantage, but sometimes it's the reason why a company has reached this situation in the first place. Can be identified by consulting the decision-maker.
Value of a choice	Every choice in a decision is related to a specific value. These values are important to take into account before deciding. Identifying these values requires lots of research, and most of the time these values become clear only after the decision has been made.
Gather relevant info	As much information as possible needs to be collected before being able to make a good decision.
Nr. of people involved in the decision	The number of people that are involved in a decision might affect the complexity of that decision. Every person has its own opinion and chances are these collide. If only a few people are involved in the decision-making, they are more likely to come to an agreement.
Decision Outcome Quality	This regard the quality of the outcome of a decision. Can only be determined after the decision had been made. Did the decision turn out to be the right one? Are the consequences and values as predicted?

Table 5: variables explained

B – Design principles

Group		Design Principles
(1) BASIC	1.1	**Basic information**
		a) Application domain and prerequisites for applicability
		b) Purpose of use
		c) Target group
		d) Class of entities under investigation
		e) Differentiation from related maturity models
		f) Design process and extent of empirical validation
	1.2	**Definition of central constructs related to maturity and maturation**
		a) Maturity and dimensions of maturity
		b) Maturity levels and maturation paths
		c) Available levels of granularity of maturation
		d) Underpinning theoretical foundations with respect to evolution and change
	1.3	**Definition of central constructs related to the application domain**
	1.4	**Target group-oriented documentation**
(2) DESCRIPTIVE	2.1	**Intersubjectively verifiable criteria for each maturity level and level of granularity**
	2.2	**Target group-oriented assessment methodology**
		a) Procedure model
		b) Advice on the assessment of criteria
		c) Advice on the adaptation and configuration of criteria
		d) Expert knowledge from previous application
(3) PRESCRIPTIVE	3.1	**Improvement measures for each maturity level and level of granularity**
	3.2	**Decision calculus for selecting improvement measures**
		a) Explication of relevant objectives
		b) Explication of relevant factors of influence
		c) Distinction between an external reporting and an internal improvement perspective
	3.3	**Target group-oriented decision methodology**
		a) Procedure model
		b) Advice on the assessment of variables
		c) Advice on the concretization and adaption of the improvement measures
		d) Advice on the adaptation and configuration of the decision calculus
		e) Expert knowledge from previous application

Table 6: Design Principles (retrieved from Pöppelbuß et al., 2011)