Universiteit Leiden
ICT in Business

Building the Data-Driven Organization:

a Maturity Model and Assessment

Name: Ruben Buitelaar

Date: August 1, 2018

1st supervisor: Niels van Weeren
2nd supervisor: Arno Knobbe

MASTER'S THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Building the Data-Driven Organization: a Maturity Model and Assessment

Ruben Buitelaar

1st August, 2018

Preface

The thesis before you, *Building the Data-Driven Organization: a Maturity Model and Assessment*, is written to fulfill the graduate requirements for the degree of Master of Science in ICT in Business at Leiden University. It was made under the supervision of Niels van Weeren.

Mobiel.nl, the company where I have been employed as a software engineer, provided me with the opportunity for an internship, in order to write my thesis on data-driven organizations. I see this thesis as an appropriate conclusion to the ICT in Business master's program. Both the program and this thesis explore how IT can not only support existing business activities, but perhaps more interestingly, how IT can drive new and exciting business opportunities.

I would like to take this opportunity to express my gratitude to everyone who has helped me in this endeavor. Foremost, my first supervisor, Niels van Weeren, who has helped me from the beginning. With his guidance, we have gone from an empty idea to a practical model, with numerous, relevant, real-world applications. I would also like to thank Arno Knobbe, my second supervisor, for taking the time to step in late into the process. His advice helped me to turn the draft into a thesis I am proud of. I express my deepest gratitude to all research participants for taking the time out of their busy schedules in order to help my research and provide me with valuable feedback. Special thanks go to everyone at Mobiel.nl, for providing me a place to work on this thesis, and for being wonderful colleagues. Finally, I thank my parents, for their love and support, which allowed me to fully focus on my studies.

Ruben Buitelaar
August 2018

Abstract

In the fast-paced business world of today, organizations struggle to capture the maximum value from their data. Data-driven organizations, organizations that excel in turning data into action, are rapidly creating a competitive advantage. Onlookers feel the urgency to join their ranks. This raises the question: How can we build a data-driven organization?

The journey to become fully data-driven is described in the novel Data-Driven Maturity Model, a framework covering known theory and best practices. The accompanying Data-Driven Maturity Assessment positions the organization on the model and shows the steps to full data-driven maturity in a personalized maturity report.

The maturity model is developed according to a published procedure for developing maturity models for IT management. An extensive literature review compares the existing landscape of business intelligence & analytics maturity models. The current literature has its limitations but offers us theory and principles to extend upon. We put the known principles in a bigger, organization-wide context.

We develop the Data-Driven Maturity Model and Maturity Assessment by iterating over multiple working versions, refining the content each time. The assessment, in the form of multiple-choice questions, is available through an online application at `https://data-driven.rubenbuitelaar.com`. An assessor can finish the assessment in 30 minutes, after which they are presented a Data-Driven Maturity Report with a maturity overview and relevant action points. Attached to the assessment is a meta-evaluation about the model and assessment. This meta-evaluation, conducted at the same time as the maturity assessment, is used to validate and refine the maturity model and assessment.

The results of the meta-evaluation validate the content and usefulness of the developed maturity model. Five stages of progress: Reporting, Analyzing, Optimizing, Empowering, and Innovating. Ten dimensions of data-driven maturity: Data, Metrics, Skills, Technology, Culture, Leadership, Strategy, Agility, Integration, and Empowerment. These stages and dimensions cover the most important principles of data-driven maturity. The maturity assessment is found to be both accurate and practical.

The outcome of the assessments shows us that many organizations are struggling to mature beyond basic analytical capabilities and find themselves in the beginning stages. Analysis suggests that organizations might have a stronger focus on the more technological factors rather than organizational factors. With the help of the Data-Driven Maturity Assessment, organizations can discover a more holistic view and use the report to create a strategic plan. Recommendations include using the maturity assessment as a tool for assessing current maturity, as a blueprint for a strategic plan, and as a measure of progress.

Contents

1 Introduction **5**
1.1 Background . . . 5
1.2 Evidence-Based Decision-Making . . . 6
1.3 Problem Statement . . . 7
1.4 Scope & Context . . . 7
1.5 Research Questions . . . 8

2 Methodology & Structure **11**
2.1 Designing Maturity Models . . . 11
2.2 Maturity Framework Development . . . 12
2.3 Literature Review . . . 13
2.4 Maturity Assessment . . . 14
2.5 Data Analysis . . . 15

3 Literature Review **16**
3.1 Research Protocol . . . 17
3.1.1 Search Strategy . . . 17
3.1.2 Study Selection . . . 17
3.1.3 Study Quality Assessment . . . 17
3.1.4 Data Extraction . . . 18
3.2 Review of Maturity Models . . . 19
3.2.1 TDWI Analytics Intelligence Maturity Model . . . 19
3.2.2 Maturity Model for Business Intelligence and Performance Management . 21
3.2.3 The HP Business Intelligence Maturity Model . . . 22
3.2.4 Big Data Business Model Maturity Chart . . . 23
3.2.5 Data Science Maturity Model . . . 24
3.2.6 Big Data & Analytics Maturity Model . . . 25
3.2.7 Adobe Analytics Maturity Model . . . 27
3.2.8 The Five Stages of Analytical Maturity . . . 28
3.3 Conclusions . . . 30
3.3.1 Grey Literature . . . 30
3.3.2 Trends . . . 30
3.3.3 Knowledge-Doing Gap . . . 31
3.3.4 Summary . . . 31

4 Maturity Model Composition **32**
4.1 Design Strategy . . . 32
4.2 Composing Dimensions . . . 32
4.2.1 Beyond Analytics . . . 34

4.3 Composing Stages . . . 34

5 Maturity Assessment **36**
5.1 Purpose . . . 36
5.2 Design . . . 36
5.2.1 Quantitative versus Qualitative . . . 37
5.2.2 Question Types . . . 37
5.2.3 Collection Method . . . 38
5.2.4 Scoring Model . . . 38
5.2.5 Evaluation Standards . . . 38
5.2.6 Questionnaire . . . 44
5.3 Meta-Evaluation . . . 51
5.3.1 General information . . . 52
5.3.2 Stages . . . 52
5.3.3 Dimensions . . . 52
5.3.4 Attributes . . . 52
5.3.5 Model, Results & Report . . . 53
5.4 Report . . . 53

6 Iterative Development **55**
6.1 Strategy . . . 55
6.2 First phase . . . 56
6.3 Second Phase . . . 57
6.4 Third Phase . . . 58

7 Final Maturity Model **59**
7.1 Dimensions . . . 59
7.1.1 Data . . . 59
7.1.2 Metrics . . . 60
7.1.3 Skills . . . 61
7.1.4 Technology . . . 61
7.1.5 Leadership . . . 63
7.1.6 Culture . . . 63
7.1.7 Strategy . . . 64
7.1.8 Agility . . . 64
7.1.9 Integration . . . 65
7.1.10 Empowerment . . . 65
7.2 Stages . . . 66
7.2.1 Reporting . . . 66
7.2.2 Analyzing . . . 69
7.2.3 Optimizing . . . 73
7.2.4 Empowering . . . 76
7.2.5 Innovating . . . 80

8 Results & Discussion **84**
8.1 Data-Driven Maturity Model . . . 84
8.1.1 Stages . . . 84
8.1.2 Dimensions . . . 86
8.2 Data-Driven Maturity Assessment . . . 89
8.2.1 Validity . . . 89

8.2.2 Accessibility . . . 92
8.2.3 Accuracy . . . 95
8.2.4 Viability . . . 96
8.2.5 Usability . . . 98
8.3 The State of Data-Driven Maturity . . . 98
8.3.1 From Data Science to Data-Driven . . . 100
8.4 Research Limitations . . . 101
8.4.1 Sample Selection . . . 101
8.4.2 No Longitudinal Observations . . . 102
8.5 Further Discussion & Additional Observations . . . 102
8.5.1 Data Dimension . . . 102
8.5.2 Local versus Global Scope . . . 102
8.5.3 Theory Formalization . . . 102

9 Recommendations & Applications **103**
9.1 Introduction . . . 103
9.2 As a Standalone Maturity Assessment . . . 103
9.3 As a Data-Driven Maturity Framework & Blueprint . . . 104
9.4 As a Cross-Sectional Assessment . . . 104
9.5 As a Measure of Progress . . . 104
9.6 Summary . . . 104

10 Conclusions & Future Work **105**
10.1 What is a data-driven organization? . . . 105
10.2 What does a data-driven maturity model look like? . . . 106
10.3 What does a data-driven maturity assessment look like? . . . 107
10.4 Research Relevance . . . 108
10.5 Future Work . . . 109

11 Appendices **114**
11.1 Example Report . . . 114
11.2 Assessment Application Homepage . . . 121

List of Figures

1.1 From classical to data-driven organization. . . . 7
1.2 From an as-is situation to a to-be situation. . . . 9
1.3 Research components. . . . 9

4.1 Identified model dimensions 33

6.1 Iterative development strategy. . . . 55

8.1 Stage importance. . . . 86
8.2 Dimension importance. . . . 88
8.3 The accuracy of report scores for all dimensions and overall score. . . . 96
8.4 The survey response on intent, strategy, information, purpose, scope, and results 97
8.5 Average maturity score and standard deviation for every dimension. . . . 99
8.6 Distribution of overall maturity scores. . . . 100

11.1 Homepage of the assessment application. . . . 121

List of Tables

8.1 Chi-squared test for survey response on stage purpose for all stages. . . . 85
8.2 t-test for survey response on stage importance for all stages. . . . 85
8.3 Chi-squared test for survey response on dimension purpose for all dimensions. . . 87
8.4 t-test for survey response on dimension importance for all dimensions. . . . 88
8.5 t-test for survey response on attribute importance for Data, Metrics, Skills, Technology, and Leadership. . . . 90
8.6 t-test for survey response on attribute importance for Culture, Strategy, Agility, Integration, and Empowerment. . . . 91
8.7 Chi-squared test for survey response on attribute knowledge for Data, Metrics, Skills, Technology, and Leadership. . . . 93
8.8 Chi-squared test for survey response on attribute knowledge for Culture, Strategy, Agility, Integration, and Empowerment. . . . 94
8.9 t-test for survey response on report accuracy for all dimensions and overall . . . 95
8.10 Chi-squared test for survey response on purpose, scope, and results 97

Chapter 1

Introduction

In today's increasingly digitized world, data is the common denominator. Leading organizations are turning data into valuable insight and powerful capabilities. While these leading organizations increasingly rely on data for their decision-making process, others struggle to capture value from data and fail to fulfill this data-driven dream [1].

Organizations feel an increasing urgency to adapt, in order to not fall behind the competition. However, they often lack the knowledge to change their old, intuition-based, way of working. With the right knowledge and tools, these organizations can also become successful in a world of data, and achieve full *data-driven maturity*.

The urgency to transform has spawned a proliferation of grey literature in a range of related domains. Although the produced knowledge can be very helpful, it has no formal design methodology, lacks public documentation, and is too often derived from anecdotal evidence. There is a need for formal academic documentation with a clear methodology and model validation.

In this thesis, we develop a *Data-Driven Maturity Model*, a framework that covers the known theory and best practices of data and business analytics. With an accompanying *Data-Driven Maturity Assessment*, we help organizations to assess their current capabilities and show the steps toward a true *data-driven organization*.

1.1 Background

The idea and desire for a data-driven organization did not come out of nowhere. Organizations have been doing data-driven activities for a long time to pursue more accurate and effective decision-making. The advent of the computer set off research efforts in ways computers could assist business processes by quickly turning information into business value [2]. During the 1960's, Michael S. Scott Morton pioneered *decision support systems (DSS)*, which assisted managers in making recurring decisions [3].

Hans Peter Luhn introduced the term *business intelligence (BI)*, proposing techniques to extract information from documents in order to provide intelligence for certain processes within an organization [4]. Luhn adapted his definition of *intelligence* from the Webster's dictionary: "the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal" [4]. We can define BI as all activities that focus on 'enriching the facts' to create information that is more actionable than raw data. These activities can range from visualizing related facts to capturing advanced insight over complex patterns.

Business analytics (BA) is a term that has been gaining popularity in recent years. BA and BI are often used interchangeably by practitioners. The origins of analytics can be found in management exercises in the late 19th century, long before the first computers were available [5]. Analytics referred to the process of collecting and analyzing business information in order to optimize the business processes. These days, BA usually refers to the knowledge discovery component of BI, while BI comprises a broader scope such as reporting and visualizing data [6]. BI and BA are often grouped together as *business intelligence & analytics (BI&A)* or just *analytics.*

The rise of the internet ushered in a new era for analytics, with new challenges and possibilities. Increased network speeds and storage capabilities meant the amount of data available for collection and analysis increased rapidly. *Big data* is used to describe data of such large volume that it cannot be processed by conventional means.

All of the mentioned capabilities are stand-alone activities. They are often isolated in a separate role or department. Nowadays, more business processes are becoming digitally integrated. Data flows everywhere throughout the entire organization. Being data-driven as an organization means supporting your decisions with data-backed intelligence. But being data-driven is also about transcending isolation and integrating data-driven activities into your business processes. The goal is to enable all employees, not just business analysts or data scientists, to explore and exploit data. Data-driven organizations are those who have successfully empowered employees with data-driven capabilities: enabling them to optimize and innovate.

1.2 Evidence-Based Decision-Making

An organization is a group with common objectives. To reach these objectives, they have to make the right decisions. Understanding the process of decision-making in organizations has been of great interest to many researchers for a long time. Decisions based on the analysis of the available information leads to evidence supporting a decision. Often, there is no time or ability to follow this process, so decisions within organizations become based on intuition or the opinion of a high-ranking employee [7].

Humans are often irrational decision-makers. In the 1950's, Herbert Simon introduced *bounded rationality*, the theory that individuals are limited in their ability to process information [8]. This limits the rationality in their decision-making process and leads to suboptimal decisions. Simon's work in artificial intelligence also showed that computers could play an important role as decision support systems because of their information processing abilities [9].

We can see that organizations do not make rational decisions. Studies show that a rational decision-making process leads to better decisions and therefore better firm performance [10, 11], so there is a clear incentive to adopt a rational and evidence-based decision-making process.

A data-driven organization should strive to make their decisions based on the available data instead of intuition. The amount of data being collected is increasing exponentially, and organizations are looking for new ways to turn their data into a competitive advantage. Research has shown that companies that emphasize data-driven decision-making show higher performance [12].

1.3 Problem Statement

Why you should become a data-driven organization is often understood. Basing decisions based on evidence rather than intuition leads to better decisions and a higher performance. Yet many organizations fail to set the right steps into becoming a data-driven organization. The problem is *what to become* and *how to become data-driven*. There is a gap between intention and taking action. *What* does the future situation for a data-driven organization look like? Without knowing *what* you are heading toward you can never develop a plan to get there. How is a data-driven organization different from an organization that has a department for business intelligence & analytics?

If you do have some idea of *what* you want to become the second challenge is *how* are you going to do it. *How* are you going to transform your current situation into a data-driven organization? This is a fundamental change in the fabric of the organization. What are the pitfalls and best practices?

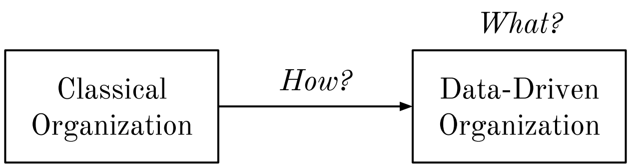

Figure 1.1: From classical to data-driven organization.

1.4 Scope & Context

The *data-driven organization* is a broad subject. Organizations from many different domains can become data-driven. We therefore not scope our research on a specific domain (e.g. healthcare or e-commerce), instead, we aim to formulate common principles from data-driven organizations in different domains. Data-driven organizations rely on digitization, so we only include organizations that are for a large part digitized themselves. The size of the organization has a large influence on how data-driven will manifest itself. Small organizations may not have dedicated departments or roles that focus on analytics. Very large organizations may require a more customized approach and may not benefit from a generalized data-driven maturity model. Our maturity model will attempt to not be opinionated toward a certain organization size, although examples may be better applicable for specific organizations. We also have to recognize that very small organizations (1 to 20 employees) may not have the resources available to invest in data-driven activities.

To summarize, our core audience is organizations with the following characteristics:

- With 50 to 5.000 employees.
- In a digitized domain (e.g. e-commerce, finance).
- With a desire to become data-driven.

We aim to provide a holistic view on being data-driven. This means we will not focus on specific activities of a data-driven organization, such as data-driven marketing or machine learning. Our aim is to study the foundation of the data-driven organization and only provide examples of what can be built on top of that. A holistic view also means that we take into account the underlying factors like culture and empowerment instead of merely focusing on analytical capabilities.

The context of this research is important to note because it may bias our views. Many of the best practices discussed in this thesis are based on American of European organization archetypes. Many of the analyzed maturity models are of American origin. Our validation is done by European (Dutch) participants. Cultural differences may lead to biases. One culture might be more receptive to strong leadership while another culture might be more receptive for change from the bottom up.

The background of the researcher is in e-commerce. Although the maturity model will aim to be agnostic about the application domain, best practices and examples may be better suited for certain application domains.

Our background context:

- Background of organizational theory: American & European.
- Background of researched maturity models: Mainly American.
- Background of the researcher: European (Dutch); background domain: e-commerce.
- Background of research participants: European (Dutch).

1.5 Research Questions

The main challenge is that organizations struggle to become data-driven. They both lack the knowledge to determine their direction, *what are we aiming for?*, and to determine a change process, *how are we going to do it?*

In change management, a key understanding is to start with where you are and to establish where you want to go. From an *as-is situation* toward a *to-be situation*. If you have the beginning and the end you can formulate a strategic plan for change. Our aim is to help organizations in this complicated process by providing the right tools and information.

With a maturity model, we can establish a desired to-be situation based on known experience and best practices. The maturity assessment positions an organization on the maturity model and determines the as-is situation of the organization. When we have established the current state and the future direction we can provide relevant information about the current challenges and opportunities. A maturity report provides what can be the fundamentals of a strategic plan to become a mature data-driven organization.

Figure 1.2: From an as-is situation to a to-be situation.

The overall objective of this thesis is to help organizations become data-driven by providing the right set of knowledge and tools. We divide this objective into three research components: *Theory*, *Artifact*, and *Practice*.

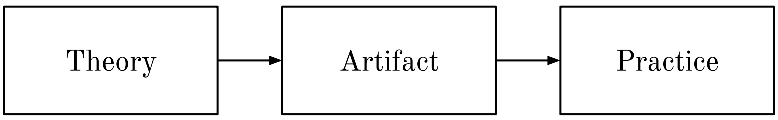

Figure 1.3: Research components.

1. **Theory**
   The first objective is to establish what data-driven means and what a data-driven organization looks like. What are the characteristics and benefits? What would an ideal data-driven organization look like? Having a reference point as a potential to-be situation helps in creating a plan for how to reach that desired situation.

2. **Artifact**
   To be able to apply the theory, we create a design artifact according to guidelines set by formal design methodologies. A maturity model that we can use as a bridge to bring theory into practice. We use the established theory to design an artifact that describes the journey of a classical organization to a data-driven organization accompanied by challenges, opportunities, and best practices.

3. **Practice**
   The last objective is to bring the artifact to practice. We do this by creating a data-driven maturity assessment that positions the organization on the maturity model and provides the most relevant information. The assessment provides an as-is situation: a starting point for change.

To fulfill these research objectives, we must answer the following research questions:

1. What is a data-driven organization?
   (a) What is the difference between business intelligence & analytics and data-driven?
   (b) What are the characteristics of a data-driven organization?
2. What does a data-driven maturity model look like?
   (a) What stages does a data-driven maturity model have?
   (b) What dimensions does a data-driven maturity model have?
3. What does a data-driven maturity assessment look like?
   (a) What questions does a data-driven maturity assessment have?
   (b) Can the outcome of the data-driven maturity assessment help organizations in settings steps toward data-driven maturity?

Chapter 2

Methodology & Structure

In this chapter, we will describe the general structure of the thesis and research methodology used to answer the research questions. We also discuss why these methodologies are suitable for our research. Our hypothesis is that a data-driven maturity framework, consisting of a maturity model and assessment, can help organizations become more data-driven. To develop the maturity model, we will use a known procedure that will provide the general structure of the thesis. We will also discuss additional research methodologies we will use.

2.1 Designing Maturity Models

Maturity models have been used extensively to guide organizations toward domain maturity. However, the concept has received its share of criticism. We discuss some main points of criticism raised by Lahrmann et al. [13] and how we aim to avoid these issues.

- **Theoretical Foundation**
  The lack of a formal theoretical foundation is an often heard criticism of the Capability Maturity Model and maturity models in general [14, 13, 15]. Maturity models are often emerging from practice and rely on anecdotal evidence and best practices. A lack of result validation deepens this sentiment. We aim to combat this by validating the maturity model with industry practitioners to test for effectiveness and usability.

- **Design Methodology**
  The lack of formal design methodologies used in the construction of the maturity model is also an often raised issue [14]. More recently design methodologies for maturity models based on the principles of *design science* have been published. We are going to use one of these design methodologies aimed at developing maturity models for IT management.

- **Situational Factors**
  Mettler et al. [16] pose that maturity models often fail to take into account the behavioral field of organizational theory and lack contextual factors such as organization size. Emphasis on capabilities and process might undermine the important influence of contextual factors such as culture and people's capabilities. The emphasis of process over people is one of the most heard criticisms of the CMM [15].

  In our maturity model design, we aim to provide a holistic view of data-driven factors within an organization. In our view, people and culture are an important factor in becoming a data-driven organization. We will include these factors in our model and test if practitioners find them critical to data-driven success.

Other contextual factors, such as organization size, are harder to fit in. Every organization is unique and we aim to provide principles that are applicable to all organizations. We will attempt to validate if organization size has an effect on whether certain factors are deemed less or more relevant.

- **Knowledge-Doing Gap**
  Pfeffer et al. argue that the purpose of maturity models is to identify the gap between knowledge and practice. Yet a lot of models fail to describe how this gap should be closed [17]. Documenting how to close the gap between knowledge and action is a challenge because every situation is unique. However, driving organizations to action is one of the goals of our maturity model. So we aim to provide advice that is specific yet agnostic about the organizational context. We will test how practitioners feel the model closes the knowledge-doing gap.

- **Documentation**
  A lack of documentation on the maturity model prevents peers from reviewing the methodology and process behind the model creation [13]. We should not have to rely on good faith to accept a maturity model. The design process of this maturity model will be fully documented in this thesis and will be publicly available.

2.2 Maturity Framework Development

In the paper *Developing Maturity Models for IT Management – A Procedure Model and its Application* [18], Becker et al. lay out a procedure for developing maturity models based on the guidelines set in *Design Science in Information Systems Research* [19]. These standards and procedures help in developing the maturity model and provide a general framework for the structure of this thesis.

The following phases are included in the procedure model:

1. **Problem Definition**
   Defining the problem, domain, audience, and scope.

   The *problem definition* is expanded in the *Introduction* chapter. We also addressed the domain, audience, and scope of the maturity model.

2. **Comparison of Existing Maturity Models**
   A comparison with existing maturity models to examine the existing literature.

   Chapter 3: Literature Review will consist of a literature review of existing maturity models in related domains. We will use these maturity models as a foundation of our own maturity model.

3. **Determination of Design Strategy**
   Decide which design strategy to use. Basic design strategies are ‘completely new model design’, ‘enhancement of an existing model’, ‘combination of several models into a new one’, or ‘transfer of structures or contents from existing models to new application domains’.

In *Chapter 4: Maturity Model Composition*, we develop the initial iteration of the maturity model. This chapter starts with the *determination of design strategy*.

4. **Iterative Maturity Model Development**
   Iterative development of the maturity model in the following sub-steps: *selecting the design level*, *selecting the approach*, *designing the model section*, and *testing the results*.

   We will use a slightly different procedure from the specified procedure model. The first iteration is developed in *Chapter 4: Maturity Model Composition*. We then develop the maturity assessment and implement the transfer of media so we can use the results to create new iterations of the maturity model.

   In *Chapter 6: Iterative Development*, we further refine the model iteratively using the results of the evaluation.

   We have to be careful that the participants in our initial iterations are representative of our audience. If we have a biased sample we may refine the model for an audience that is not representative of our intended audience.

5. **Conception of Transfer and Evaluation**
   Determine how we communicate and transfer the maturity model to the target audience.

   In *Chapter 5: Maturity Assessment*, we develop the maturity assessment that accompanies the maturity model. In this chapter, we also develop an evaluation method to determine how well the maturity model works. We also explain the transfer method.

6. **Implementation of Transfer Media**
   Make the maturity model accessible to the target audience.

   How we implemented the transfer of media is explained in *Chapter 5: Maturity Assessment*

7. **Evaluation**
   Establish if the model provides the expected benefits.

   We will determine the benefits of the model in *Chapter 6: Iterative Development*. We use these results to refine the maturity model. After development, we evaluate the model in *Chapter 8: Results & Discussion*.

8. **Rejection of Maturity Model**
   In the case of negative evaluation results, the model can either be further adjusted into a new version or be taken off the market.

   In *Chapter 10: Conclusion*, we will conclude whether the model provides the expected benefits and therefore should continue to be available or taken off the market.

2.3 Literature Review

The literature review in Chapter 3 is conducted to establish an initial version of the maturity model and find relevant information within existing research. The main part of the literature review consists of a *systematic review* of existing maturity models relevant to our research domain. We will use the guidelines for systematic review as specified in *Procedures for Performing Systematic Reviews* [20]. Systematic review has its origins in the medical research. In *Procedures for Performing Systematic Reviews*, Kitchenham has adopted the systematic review practices to the field of software engineering. We will use these steps and guidelines in our systematic review of existing maturity models. Systematic review is used to clearly define the process of a literature review. If the process is strictly defined and followed, the review can be assessed and reproduced.

The process defined by Kitchenham defines the following steps for planning and conducting the systematic review:

1. **Review Objectives**
   Identify the need and objective for a systematic review.
2. **Review Protocol**
   Develop a review protocol where you pre-define the methods you are going to use in the systematic review.
3. **Search Strategy**
   Execute a predefined search strategy to systematically find studies.
4. **Study Selection**
   Assess studies for relevance.
5. **Study Quality Assessment**
   Assess relevant studies on predefined quality criteria.
6. **Data Extraction**
   Extract predefined information from the studies.

2.4 Maturity Assessment

In this section, we will discuss our methodology for assessing data-driven maturity.

The assessment consists of two parts: The assessment part with questions that will position the organization on the maturity model and a meta-evaluation that is used to evaluate if the maturity model and assessment deliver the projected benefits.

Our methodology for designing the maturity assessment is based on the paper *Meta-assessment: Evaluating Assessment Activities* by John C. Ory [21]. This paper discusses standards set by the *Joint Committee on Standards for Educational Evaluation* [22] on how to evaluate assessments used in education.

Although the subject and domain are different from our own assessment the discussed design standards help us in our initial design of both the assessment as our own meta-evaluation. The standards are already being used beyond their original domain, but we have to be careful not to misapply the standards [23]. In our situation, we have to be aware of the fact the standards are originally designed to evaluate people, while we are evaluating organizations. Four types of standards are discussed. Utility standards ensure 'the practical information needs of given audiences'. Feasibility standards ensure 'the assessment will be realistic, prudent, diplomatic, and frugal'. Propriety standards ensure the assessment will be conducted legally and ethically. Finally, accuracy standards ensure assessment convey adequate information about the studied object. In *Section 5.2: Maturity Assessment Design* we discuss how we apply these standards in the design of our maturity assessment.

2.5 Data Analysis

The outcome of the survey on the maturity assessment will be analyzed using statistical procedures. In addition to qualitative feedback boxes, we have two types of quantitative questions in our survey: a Likert-type question and scalar-type question. The scalar goes from 0 to 10. For our scalar questions, we will use methods usual for continuous values. We will use a *Student's t-test* to determine if the response is significantly different from a neutral normal distribution around 5. The Student's t-test is used because we do not know the population variance, which we have to estimate from our sample. The Z-test was not used because we have a low sample size ($n < 30$).

The statistical analysis of a Likert-type question is subject to great discussion. Many argue that the steps on a Likert scale do not represent a continuous measurement [24]. The step from *agree* to *strongly agree* does not represent the same distance as *agree* to *neutral*. Therefore, Likert scale questions produce ordinal data. We will use the *chi-squared test* to analyze this data. We will group the negative responses, *strongly disagree* and *disagree*, and the positive responses, *agree* and *strongly agree*. The neutral responses, that neither agree or disagree, are ignored in the test. We can then determine if the response is significantly different from a neutral response, which would be half negative and half positive. We use an alpha level of 0.05 for all our statistical tests.

Chapter 3

Literature Review

To create a strategic plan for change, it is beneficial to start with where you are. Only that way you can formulate the change necessary to reach the desired state. To establish an *as-is* state, you can use a frame of reference in the form of a *maturity model*.

A maturity model is a model that is composed of different stages of maturity for a certain discipline. Maturity refers to the ability of the organization to consistently deliver the intended benefits of the domain. The *Capability Maturity Model (CMM)* was first introduced in 1988 by Watts Humphrey for assessing the capabilities of software development organizations [25]. It provides the ability to both assess the current maturity of software development within an organization and act as guidance in achieving a more mature process. The CMM defined five levels of process maturity: *initial*, *repeatable*, *defined*, *managed*, and *optimized*. The goal is to bring the software development process from an initial state, where there is no statistical process control, to a state with strict process control wherein the process can be optimized.

Inspired by the CMM many maturity models emerged in related application domains. Maturity models are now found in disciplines such as quality management, project management, business process management, information security management, and analytics. A maturity model provides a convenient way to assess the current situation of an organization and provides guidance for continually improving the relevant processes. Typically they start with an initial stage where the activities in the domain have just been initiated. The stages progress to a level where the capabilities in the domain are fully matured. This level is considered an ideal or desired state by the creators of the model. Between these stages is a path of continuous improvement.

Many maturity models are created and used by consultancy firms, leading to many competing maturity models with each their own nuances and focus areas. Some models have accompanying methods to evaluate the current state of the organization. This can be done with for example questionnaires, in which employees have to indicate on a Likert scale how the organization scores on specific questions. The results of the questionnaire are used to present an as-is situation.

Other models rely on qualitatively comparing the current situation with the described stages. This can be done with relative ease for technological aspects, such as which analytics techniques are being used. It is harder to compare 'soft' aspects of the organization, such the culture of the organization. Models can provide examples of the culture at a particular stage, which can be used for comparison, but it is harder to effectively measure the progress an organization is making.

3.1 Research Protocol

In this section, we discuss our research protocol for the systematic review of maturity models within our domain. We define what we are looking for, how we are going to find it, and how we are going to process it.

3.1.1 Search Strategy

In order to create a data-driven maturity model, we first have to analyze the existing maturity models in the related application domains. We will evaluate their strengths and weaknesses to see which aspects we can incorporate into a data-driven maturity model.

Selection of models was done by searching for maturity models in the following domains: *business intelligence*, *business analytics*, *big data*, *data science*, and *data-driven marketing*. We also looked at meta-analyses of BI maturity models as a starting point for finding more suitable maturity models.

3.1.2 Study Selection

The goal is not to do a comprehensive analysis of all the relevant models that are available. There are too many models to analyze, and some are only minor variations of existing models. The aim is to make a selection of models that cover all the relevant domains and dominant design perspectives. We have applied some criteria to determine which maturity models we analyzed.

The model has to be publicly documented. Some commercial models do not provide sufficient documentation to analyze, but only provide a questionnaire to (self-)assess the organization. Public documentation allows the model to be analyzed for its strengths and weaknesses and compared to other maturity models. Consultancy organizations often do not publicly give out the associated advice for a given stage because this is part of their consultancy services.

The model should reflect recent developments. There have been a lot of recent (technological) developments in the field of analytics and big data. Useful maturity models should reflect these changes. Technology changes the way the organization tackles analytics. Recent developments, such as the surge in big data, change the way an organization governs their analytics processes. Many older models, such as the Data Warehouse Maturity Model [26], used to focus on one specific technology that was relevant at the time. The model evolved over time and has led to TDWI's Business Intelligence Maturity Model [27].

3.1.3 Study Quality Assessment

Chosen studies have to be assessed for their quality. This is a challenge for our literature review because we want to include gray literature. This literature has not been published and often lacks comprehensive documentation and methodologies, which is also one of the main criticisms at maturity models in general [16]. However, the nature of this domain is for a large part focused on commercial applications, and many best practices are therefore generated outside the scientific community. Our aim is to formalize these practices and be extremely aware of the origin of the maturity model.

3.1.4 Data Extraction

Many maturity models have been created in the field of analytics. With the help of a maturity model, organizations can assess their current capability and learn from the best practices in analytics. In order to analyze the maturity models, we first have to define some common characteristics we can extract for comparison.

Domain

The first and foremost difference between maturity models is the application domain the model is intended for. This domain dictates for a large part the direction, focus, and scope of the model. A model for big data has a large focus on the maturity of the big data processing capabilities. A model that covers business analytics might focus on predictive and prescriptive analytics and also cover big data related capabilities. There are many different domains related to data-driven for which maturity models have been developed, of which many domains are very closely related. We have found models in the following data-driven domains for our analysis: *business intelligence*, *business analytics*, *big data*, and *data science*. The domains might overlap a lot but we have to be aware of the direction and goal the model has.

Dimensions

Organizations are often evaluated within a domain on one or more of the following dimensions: *process maturity*, *object maturity*, or *people capability* [28]. Process maturity refers to 'how explicitly a process is defined, managed, measured, controlled, and effective'. Object maturity is 'the predefined level of sophistication a particular object has reached'. This could be for example a software product or data warehouse environment. For analytics, we can also speak of *technological maturity*. People capability is 'the ability of the workforce to enable knowledge creation and enhance proficiency'. Many maturity models enhance or add new dimensions that are relevant for the domain. For example, the TDWI Analytics Intelligence Maturity Model has the following five dimensions: *Organization* (people), *Infrastructure* (technological), *Data management* (process), *Analytics* (technological), and *Governance* (process) [29]. Some models only focus on a specific dimension within the domain. The original CMM is criticized for its focus on the process of software engineering while disregarding the people [15]. We can both look at the different dimensions the model has, and which dimensions have more focus.

Stages

Maturity models are usually divided into stages of maturity. Although the process of maturing is a continuous process, the division into stages has its advantages. The maturity can more easily be assessed if you couple distinct capabilities and scenarios to a certain stage. Instead of aiming for a very accurate assessment on a continuous spectrum, you can now place the organization within a stage. Specific guidelines and advice can then be provided for the stage, so it becomes clear where the current priorities lie. Stages are also easier to visualize and communicate within the organization.

We see a lot of difference between the design of the stages. Some models opt for stages focused on distinct technological capabilities, while others focus on the organizational state. This also depends on the dimension focus of the maturity model.

Realism versus Idealism

We can make another distinction in to which extent a maturity model illustrates a realistic scenario. A realistic model opposes an idealistic model. A more realistic model presents the stages as scenarios you as an organization will typically go through, acknowledging the immaturity and inefficiency in early stages. Realistic models often have a lot of focus on the shortcomings the organization has to resolve to progress to the next stage. Idealistic models have a stronger focus on the specific capabilities you have to build for progression. Idealistic models often have an emphasis on the technological dimension without addressing the organizational challenges that have to be overcome.

Publication type

Many maturity models are published as a form of thought leadership, to prove the creators have expansive knowledge on the subject. The maturity models do not always have a scientific background and are not peer-reviewed or validated. We have to be aware of this and take into account how the model was published. We will look for the type of publication and the extensiveness of the documentation.

3.2 Review of Maturity Models

3.2.1 TDWI Analytics Intelligence Maturity Model

	General information
Creator	TDWI Research
Focus area	Analytics & big data
Year	2014-2015
Publication	white paper [29]
	Dimensions
Organization	Organizational factors that facilitate analytics success, such as strategy, culture, leadership, skills, and funding.
Infrastructure	How advanced is the technological infrastructure in place to support analytics?
Data management	How does the organization manage its data in support of analytics, factors such as data quality, data processing, data integration, and data access.
Analytics	How far advanced is the analytics used within the company?
Governance	How well does the data governance strategy support its users' data discovery and analytical explorations?
	Stages
Nascent	There is no official analytics process, but there may be some interest. The culture is not analytic, and decisions are based on intuition.

Pre-Adoption	Interest in analytics is growing, and research efforts are being made. Employees and leadership are starting to see the benefits of analytics and are looking for ways to start up efforts.
Early Adoption	The organization is starting to put analytics tools and processes in place. Analytics becomes a part of the decision-making process.
The Chasm	After initial success, organizations face a series of challenges before maturing to corporate-wide adoption. Advanced analytics requires more organization and more funding. It can be hard to make this happen because of responsibility issues.
Corporate Adoption	Organizations entering the corporate adoption phase have addressed the challenges and change the way business is done by incorporating analytics in their daily processes.
Mature/Visionary	Few companies can be considered visionary in analytics. These companies have streamlined the analytics processes and source their data from everywhere. The culture is highly analytical and they are continually improving their analytics processes.

The *TDWI Analytics Intelligence Maturity Model* [29] is a widely used model for assessing analytics maturity. It has an organization-centric approach with less focus on technology. The model has its roots in *TDWI's Business Intelligence Maturity Model*, which has stages based on the development of an infant to a mature adult [27].

The model's dimensions cover many aspects. The data management is a clear addition to the usual dimensions, this is a dimension we see more often in analytics maturity models. Data management mentions "data quality and processing as well as data integration and access issues". This seems to describe data governance. However, the model also contains a separate Governance dimension which describes the company's data governance strategy. Data governance is divided into two dimensions: data quality and integration versus data access and stewardship. Analytics and Infrastructure are both technological dimensions with their own focus. Analytics focuses on the application layer and infrastructure on the infrastructure and database layer of the IT. The Analytics dimension also takes into account the skills of the employees in analytics and data science. The Organization dimension is a valuable addition because it shows the important role of organizational factors such as culture and sponsorship, which are essential for the successful implementation of analytics.

The model realistically describes the stages of organizational maturity a company would go through. This is very helpful because you then know what struggles to expect and hopefully avoid. The naming of the stages is unfortunately not very descriptive. It hard to gauge what level of capability an organization has at any stage because there is not much information about the technological level. The Chasm stage is a special stage that serves to warn for the challenges an organization faces when maturing. The challenges: skills of employees, cultural and political issues, and governance are often preventing companies from reaching corporate adoption. Being aware that these issues may arise, combined with the right advice, can help a tremendous amount in overcoming them. The stages of the model are great in describing the current state of the organization and solving today's problems but might be not as useful in setting goals for growing maturity. For example, Early Adoption is not a state a company would strive for if you are currently Nascent. You can only strive for Corporate Adoption or Visionary.

3.2.2 Maturity Model for Business Intelligence and Performance Management

	General information
Creator	Gartner
Focus area	Business intelligence & performance management
Year	2010
Publication	Research report [30]
	Dimensions
Business drivers	Are the drivers for success of the program in place?
People	Do the employees have the right mindset and skills?
Program management	How mature is the management of the program?
Processes	How defined are the processes?
Tools	How advanced are the used tools?
	Stages
Unaware	Ad hoc use of BI. No formal processes, practices, and performance metrics are in place and there is extensive use of spreadsheets for analytical needs.
Opportunistic	BI and analytics projects are undertaken individually to optimize a single process. Multiple applications proliferate across the organization guided by their own team. Value is delivered very quickly to users but the skills become stuck in silo's across the organization.
Standards	This level is marked by the increased coordination across the enterprise. Expertise is shared and technology standards start to emerge, but data and analytic models are still in silo's.
Enterprise	Enterprise strategy is guided by a framework of performance metrics. Decision processes are supported by BI applications, which are used by enterprise-wide. Models and data are integrated and accessible for all teams.
Transformative	BI becomes of strategic importance to the organization. Information is a strategic asset and BI and PM opportunities for generating revenue. The systems are set up in a way that BI applications can be rapidly integrated and users can create their own reports and dashboards.

Gartner's Maturity Model for Business Intelligence and Performance Management [30] provides an enterprise perspective to BI and PM maturity. The model has a focus on BI and its role in performance management, so there is more consideration for reporting than for predictive analytics. It addresses the important role of executive sponsorship and suggests a service-oriented approach for orchestrating the BI processes and services.

The dimensions of the model are only used for the assessment. They are not used to provide structure to the characteristics of the levels and the accompanied action points. Because of the great focus on organizational aspects, there is only a simple mention of tools for all the technology used. There is little information on the tools and techniques used.

The model's stages are suitable for assessing maturity levels. They generally follow the stages of the CMM going from unorganized to organized. The final stage is the Transformative stage, which recognizes that the role of BI is not merely in supporting business processes, but can also add value on its own. However, no examples or action points are given on how BI can help generate revenue.

The action points are a good addition to this model. Every level is accompanied with recommended actions for improvement. This provides actionable insight into the current situation which helps you move to a higher level.

3.2.3 The HP Business Intelligence Maturity Model

	General information
Creator	HP's Information Management Practice
Focus area	Business intelligence
Year	2007
Publication	white paper [31]
	Dimensions
Business Enablement	How advanced are the business needs and problems solved with BI?
Information Management	How advanced are the information solutions to serve business needs?
Strategy & Program Management	How advanced is the management as a key enabler of BI success?
	Stages
Operation: Running the business	Organizations are in the early stages of BI investments with a basic BI strategy. The business needs are focused on basic reporting and analysis. Data is sourced locally and is processed with a lot of manual effort. Projects are on a small scale and ad hoc.
Improvement: Measuring and monitoring the business	Ad hoc solution are traded for planned strategies of how to monitor business processes. Business needs still focus on reporting and analysis but there is movement toward predictive analytics. Data marts are commonplace, but there is no horizontal integration between them.
Alignment: Integration performance management and intelligence	Integration of information solutions starts to take shape which gives the ability to increase business value through complex metrics and analyses. During this stage there is a focus on governance and data management.

Empowerment: Fostering business innovation and people productivity	Information is now a powerful asset. Data management and governance are in place. Analytics has become a part of the business processes. The potential of BI is being harnessed by the front line employees and BI has become an important component of corporate strategy.
Excellence: Creating strategic agility and differentiation	The final stage is where BI has become a core part and key differentiator of the organization. Lack of information is no longer an inhibitor to strategic agility, instead the availability of information is a source of strategic agility.

HP's BI Maturity Model [31] is an older model formally focused on BI but also takes many aspects into account that today would be considered a part of analytics. Although some newer techniques such as autonomous analytics and big data are not fully represented the model still provides relevant insight.

The chosen dimensions cover business needs, information management, and strategy management. The characteristics across the stages are given for every dimension. It may be hard to distinguish between the dimensions because they are all from a management perspective. For example, how would the development of talent be classified as?

The stages are straightforward and like other models go from ad hoc solutions to streamlined BI across the organization. Additionally, there is attention for the empowerment of the front line workers, which plays an important part in creating the most value out of analytics. Self-service BI was a rising trend at the time this maturity model was created [32]. The final stage is where BI is no longer just there to support business processes but is a source of strategic agility.

There are some key takeaways from this model. In the first stage, there it is already mentioned that analytics talent is required to be successful. Lack of talent can become a major inhibitor in maturing and is not easily solved with more executive sponsorship and funding. It needs to be groomed from the start. C-level sponsorship is in this model also seen as a result of past BI success. It suggests that sponsorship for BI-projects moves bottom-up from lower level managers to higher executive management. In other models we sometimes see it the other way around: High-level sponsors drive the analytics revolution from the top. All models agree that executive sponsorship is an important success factor. Finally, this model also emphasizes the importance of right-time analytics. The most successful analytics implementations are able to deliver value faster. Real-time insight is essential for a true data-driven organization.

3.2.4 Big Data Business Model Maturity Chart

	General information
Creator	Bill Schmarzo at Dell EMC Services
Focus area	Big Data & Analytics
Year	2012
Publication	Blog [33]
	Dimensions
	None specified

	Stages
Business Monitoring	BI solutions are deployed to monitor ongoing processes. Benchmarking against previous periods and simple metrics are used to identify under- or over-performing business areas.
Business Insights	In this stage, predictive analytics is used to generate actionable insight. Instead of just reporting the data, it is refined into more actionable insight, observations, and recommendations.
Business Optimization	Analytics is embedded into the business processes to optimize the operations. In this stage, a lot of value is gained because automated systems can make optimized decisions faster and better.
Data Monetization	Analytics now becomes a part of the products and services the organization delivers. Data with insight can be sold to other parties, analytics can be integrated to make an intelligent product, or analytics can be used to upscale the customer relationship and experience.
Business Metamorphosis	The final phase is a metamorphosis of the business model into new services and products based on analytics. For example, a retailer with strong shopping optimization and recommendation can become a personal tailor service that automatically sends cloths that fit the customer's style as a subscription service.

The *Big Data Business Model Maturity Chart* [33] by Bill Schmarzo is a model specifically focused on the opportunities of big data and analytics. The model is brief and not suited for nurturing analytics maturity as it lacks comprehensive advice. However, it is a good model for showing the benefits of big data.

There is a lack of any real dimensions in this model. The model only focuses on the business model opportunities that big data and analytics offer, so there is no attention for other dimensions.

The stages of this model are focused on the business model opportunities analytics has to offer. It starts with basic monitoring and ends with a total metamorphosis of the business model. There are a couple of key insights found in the model. The Business Optimization phase is not always explicitly included in maturity models. It tells it is important to operationalize the insight gained from analytics. Only when insight flows back to the business processes you can start to optimize and integrate analytics into your products and services. So operationalizing insight is important in setting the stage to further integration and monetization. The Data Monetization phase shows us that analytics itself can become (part of) a product or service. This goes even further in the Business Metamorphosis phase where the complete business model revolves around an intelligent product or service. The role of analytics shifts from a supporting to an optimizing to a central role in the organization.

3.2.5 Data Science Maturity Model

	General information
Creator	Booz Allen Hamilton

Focus area	Data Science
Year	2015
Publication	e-book [34]
	Dimensions
	None specified
	Stages
Data Silos	In the first stage of Data Silos, there is no aggregation of data. Organization have no sense what data is available for collection and what is needed for data science activities. Data is isolated in departments and there is no central integration.
Collect	When the decision is made to create data science capabilities you start the Collect stage. The organization starts to find and aggregate data for analytics activities.
Describe	The organization has collected raw data, which is now prepared and enhanced, to be more useful for analytics activities. Some basic analytical capabilities, such as counts and search query's, are available. A lot of time is still spent on collecting data.
Discover	Hidden patterns and relationships can be found in the refined data. For example, a data scientist can make a cluster of costumers that purchase the same items.
Predict	Advanced data analysis can provide predictions for future observations. Data scientists could predict which products would perform well for a particular cluster of costumers. Data science can determine how well a certain solution to a problem would perform.
Advise	In the final Advise stage data science can not only determine how well a certain solution would perform but also which solution would perform the best.

This *Data Science Maturity Model* [34] is part of the field guide for data scientists by Booz Allen Hamilton. It is a brief maturity model providing a backdrop for the activities of data scientists. The only dimension is what data science activities are being exercised.

The stages start with the Data Silos stage. From this model, we can learn the steps required for data science. It starts with isolated data that is not actively being collected and processed. In the next stages, data is collected and refined. Sometimes we tend to think of data science as only the data analyzing activities, but we see here that the collection and preprocessing of data takes the most amount of time in the early stages of data science. We can also see three different levels of analytics: discovering opportunities, predicting the outcome of a solution to a problem or opportunity, and finding the optimal solution to a problem. Every level decreases the required human interference in going from a problem to a solution. Eventually, decisions are automated and no human interference is required anymore.

3.2.6 Big Data & Analytics Maturity Model

General information	
Creator	Chris Nott & Niall Betteridge at IBM
Focus area	Big Data & Analytics
Year	2014
Publication	Blog [35]
Dimensions	
Business Strategy	How are big data and analytics represented in the business strategy?
Information	What is the role of data and the resulting insight?
Analytics	How advanced are the analytical capabilities?
Culture & Execution	How ingrained is analytics in the culture of the organization?
Architecture	What does the information architecture look like?
Governance	How is data managed and governed?
Stages	
Ad hoc	The use of big data is discussed but not reflected in the business strategy. No architecture or governance is in place and analytics is only used to describe past events.
Foundational	The potential of analytics is acknowledged, but applications are experimental. Analytics is used to explain why things have happened. Some governance is in place.
Competitive	Business strategy encourages analytics to gain insight. Analytics is used for prediction.
Differentiating	The big data and analytics capabilities have now become a competitive advantage for the business. Predictive analytics help optimize the decision-making process. Architecture is well-defined and flexible.
Breakaway	Analytics is a source of innovation, data is a key asset. Business processes are optimized and automated. The information architecture is prepared to handle all big data challenges.

The *Big Data & Analytics Maturity Model* [35] is a simple matrix model consisting of 6 dimensions and 5 stages. The documentation is not very comprehensive; it is published in a blog post. The focus is on big data and digital engagement.

The dimensions are high-level taken from an organizational perspective. The Information dimension is interesting, as many maturity models look at raw data. Information suggests more refinement and making data more actionable. Culture is also represented and recognized as an important make-or-break factor. Business Strategy at a Breakaway level also underpins that analytics can and will lead to business model innovation. New opportunities arise that were not previously possible.

The stages in this model progress from Ad hoc to Breakaway. Ad hoc usage as a starting point for analytics is a concept we see in other models as well. It suggests starting analytical activities is from the bottom up. Local initiatives are started before an official program is put in place. ‘Breakaway’ suggests that organizations that excel in analytics can outpace the competition. Big data also has a winner-takes-all tendency, because data can fuel growth which in turn leads to more data.

3.2.7 Adobe Analytics Maturity Model

	General information
Creator	Scott Rigby & Gerardo Contreras at Adobe
Focus area	Marketing Analytics
Year	2015
Publication	Technical documentation [36]
	Dimensions
Collection	What data is collected?
Analysis	How much insight is currently gained from data?
Execution	In what manner is action taken based on analytics?
Automation	How much human interaction is automated for analytics?
Application	How does an organization tactically respond to analytical findings?
Attribution	How well does an organization credit all factors contributing to a marketing result?
Strategy	What is the level of talent development, culture, sponsorship, technology, and other processes related to analytics?
	Stages
Descriptive	Data is collected from web analytics. Analytics consists of viewing web traffic. Not enough organizational resources reserved for advanced analytics.
Diagnostic	Analysis focuses on KPI’s and understanding user behavior to drive improvements. Stakeholders can access reports when they want to. Attribution suggests recommendations for action.
Advanced Diagnostic	Data is also sourced from outbound marketing systems and profitable segments identified. Insight used to inform marketing processes. Executive sponsorship is growing.
Predictive	Outcomes of interactions are predicted and used to recommend actions in marketing systems. Stakeholders have access to real-time insight.
Prescriptive	Data includes offline and third-party sources. Interactions are prescribed to drive conversions. Analytics has a high skill level and a high influence within the organization.

The *Adobe Analytics Maturity Model* [36] focuses on digital marketing and web analytics. The model is part of a larger guide to assessing the operational readiness for implementing Adobe products. Adobe offers marketing technology products, such as a campaign manager and tools, to optimize digital advertising activities.

The chosen dimensions reflect the focus on marketing analytics. Whereas other models have a large focus on the 'people, process, and technology', this model rolls them up into the Strategy dimension. Collection and Analysis dimensions are also found in general analytics models. Execution and Application describe how organizations act on insight. Automation is an interesting dimension. A lot of analytics models focus on the manual activities of the analyst. For a data-driven maturity model, it is important to consider how the insight gained from analytics can be fed back into the business processes. The Automation dimension considers the level of insight fed back into the systems but also the speed of delivery. We can generate powerful insight in an analytics database, but if that insight is not used in time it might not be relevant anymore. This is especially key in marketing applications where you are generating knowledge about potential customers. The ability to apply generated knowledge in near real-time is a powerful capability. Attribution is a dimension especially relevant for marketing. Correct attribution allows you to measure what efforts are the most profitable. Attribution follows the data-driven philosophy that data and metrics are an important instrument to improve performance.

The stages of the model reflect the sophistication of the used analytics, as we see in more models. The higher the sophistication the less human intervention is needed between data and taking action. In addition to the level of analytics, some other trends can be seen. In later stages, more data is collected from different sources including third-party vendors. The marketing angle can also be clearly noticed by the focus on the prediction of the outcome of a user interaction, e.g. which offering would drive the most conversions. There is also a trend toward real-time monitoring & insight and self-service. At first, a report is generated and sent to stakeholders, while later on they can access real-time data themselves and do not have to wait on a report. An interesting trend in the Strategy dimension is the increasing level of influence analytics has on the overall organization. Increased recognition for and trust in analytics is a clear sign of analytics maturity.

3.2.8 The Five Stages of Analytical Maturity

	General information
Creator	Thomas H. Davenport & Jeanne G. Harris
Focus area	Analytics
Year	2017
Publication	Book [37]
	Dimensions
Data	What data do we have available and how is the quality of the data?
Enterprise	How does the strategy take into account the whole enterprise?
Leadership	How committed is the leadership to becoming an analytical competitor?
Targets	How targeted are the organization's investments toward distinctive capabilities?
Analysts	How cultivated is the analytical talent within the company?
Technology	How advanced is the technology to support (big data) analytics?

Analytical Techniques	How sophisticated are the techniques used for analytics?
	Stages
Analytically Impaired	Organizations have no analytics activities and have to already prepare for it in order to become a competitor. The organization has to fix their data quality and choose a distinctive capability as a strategic focus for analytical competition.
Localized Analytics	Analytics is now being used in some isolated business processes. It is now important to document the benefits and use that to gain executive sponsorship for bigger analytics projects.
Analytical Aspirations	Organizations now have executive sponsorship for analytics and have the aspiration to become an analytical competitor. Analytics takes a broader and more strategic perspective, starting with their distinctive capability.
Analytical Companies	The focus is now on building analytical capabilities across the organization. It is important to build a strong analytical culture while being aware of the cultural differences.
Analytical Competitors	Analytics is an important capability of the organization and has become a competitive advantage. The progress in data, processes, culture, and techniques are hard to duplicate for competitors, but companies in this stage continuously have to attempt to raise the bar.

The book *Competing on Analytics: The New Science of Winning* [37] by Davenport and Harris, released in 2007, has been a very influential book for many wishing to become strong in analytics. It explains what road companies have to take to become an *analytical competitor*, a company that uses their analytical prowess as a competitive advantage. The updated edition extended the DELTA dimensions with Technology and Analytical Techniques. The maturity model has a heavy emphasis on organizational factors for analytical maturity. It stresses the importance of executive sponsorship and the grooming of analytical talent. An important prerequisite is selecting a distinctive capability of the company on which you want to focus the analytical efforts. For some organizations, this can be the customer experience while others can excel in operations. The book outlines two approaches for becoming an analytical competitor. The first is the full steam ahead approach. Full steam ahead means the analytics has the full blessing and support of upper management, which means the localized analytics stage can be skipped. The other approach is the prove-it detour. In this approach, upper management might be sceptic of immediately starting up large-scale analytics projects. Smaller scale projects with middle management are started, which thoroughly document the benefits to garner support from upper management for larger scale projects. Davenport and Harris estimate that taking this detour can result in an extra couple of years before becoming an analytical company. The benefit is that during this time the analytical culture can be cultivated more broadly.

The dimensions of the original model together form DELTA, after the mathematical symbol for change. The original model has dimensions we see in other models, but it is interesting to see that Technology is not in the original. The T in the original model stands for Targets, the careful selecting of relevant projects which can benefit the most from analytics and are close to the distinctive capability. Technology and Analytical Techniques were added in the updated model. These dimensions become more prominent with the advent of big data. Big data presents powerful opportunities which require advanced technological and analytical capabilities.

The stages describe the journey from a company which has no analytics activities to a company that has a competitive advantage because of their strong analytical capabilities. The localized analytics stage is special because when embarking on the full steam ahead approach it can be skipped. When there is strong executive support for digital transformation there is no need for a prolonged period of testing the waters and garnering more support. This also sets up the analytical aspirations phase which is the point when the company has decided to fully embark on their analytical transformation. The only prerequisite for reaching this stage is the full support of upper management. Overall the stages are from a very high-level organizational perspective. There is no explanation of what specific (technological) capabilities need to be pursued at every stage.

3.3 Conclusions

We have reviewed a diverse set of maturity models in domains related to data-driven maturity. We can see a lot of overlap in trends between the models. In this section, we elaborate on some points found in the reviewed models which will be important in the creation of our own model.

3.3.1 Grey Literature

The first thing we can see is that a lot of models are not scientific publications. The subject is covered more heavily in grey literature, such as blogs and white papers. Grey literature has some disadvantages. The methodology for creating the maturity model is not revealed to the public. We cannot establish if the model is validated by peers. A lot of models also do not have the goal of being used for assessment but are used to quickly convey thought leadership on a subject. They have not been tested for usability and accuracy. No methods are conveyed on how to use the maturity model in practice.

There is a clear need for a formally built and validated model in the field of data-driven maturity and analytics.

3.3.2 Trends

Some trends can be spotted when comparing the different models. A lot of models start with identifying the information needs of the organization. How can we best utilize data and analytics? In the Five Stages of Analytical Maturity, this is explicitly being done by identifying the core analytical competency of the organization. Establish where you want to go first.

A lot of models ends with a state of full competency of analytics. Analytics is no longer a supporting activity but has become a core activity and differentiator of the organization. Analytical competency may even transform the business model of the organization. All models share the opinion that analytics will become a more central component of the strategy of the organization. From supporting the current organization to a differentiating role in a 2.0 version of the organization.

Almost all models explicitly establish the important role of culture and sponsorship. Support from the bottom up and support from the top down. Both are needed for the successful adoption of analytics across the organization.

Some models show the importance of empowering both employees and products with analytical competencies. Self-service BI is an important trend that aims to empower all employees with analytical powers. Products can also be equipped with analytical capabilities to create so-called 'smart products'.

Another important point is operationalizing analytical insight into processes. Automatically integrating insight can provide huge benefits by providing more relevant and faster actions. This is a major trend of big data that goes beyond classic BI capabilities. Real-time insight is continuing to play a bigger role in analytics.

3.3.3 Knowledge-Doing Gap

The reviewed models lack to activate readers or help to close the gap between knowing and doing. No information is provided on how to utilize the maturity model. There are almost no standardized assessment tools or methods. No scoring models are provided. It is hard to utilize these models on your own organization.

Gartner's Maturity Model for Business Intelligence and Performance Management provides action points. These action points help the reader to take concrete actions based on the information in the model.

3.3.4 Summary

To summarize, we want to take into account the following points when designing our model:

- We aim for a formally built and validated model, with a fully documented design process.
- Start with identifying information needs, grow to full competency, with analytics as a differentiating factor.
- Emphasizes the importance of culture and sponsorship.
- Show the importance of empowering employees and products.
- Integrate analytical insight into business processes.
- Help to close the knowledge-doing gap by providing practical applications

Chapter 4

Maturity Model Composition

4.1 Design Strategy

According to Becker et al. we first have to choose an overall design strategy. Basic design strategies are 'completely new model design', 'enhancement of an existing model', 'combination of several models into a new one', or 'transfer of structures or contents from existing models to new application domains'.

We have conducted a thorough literature review of existing maturity models in related domains. We choose to combine the knowledge of these maturity models and, where applicable, add new knowledge into a new maturity model. We do this by analyzing existing model dimensions and stages to compose a comprehensive maturity model. This model can then be further iteratively refined where necessary.

4.2 Composing Dimensions

The first thing to notice is the more than 30 dimensions that can be found in the analyzed maturity models. Many dimensions (partially) overlap or use different terminology. To derive useful dimensions for our model, we first have to categorize and find commonalities between dimensions. We also look if some dimensions are a subset of a larger dimension which it can be included in.

$$\text{Leadership}\left\{\begin{array}{l}\text{Enterprise}\\\text{Leadership}\\\text{Organization}\end{array}\right.\qquad\left.\begin{array}{r}\text{Collection}\\\text{Data}\\\text{Information}\\\text{Data Management}\\\text{Governance}\end{array}\right\}\text{Data}$$

$$\text{Culture}\left\{\begin{array}{l}\text{Culture \& Execution}\end{array}\right.\qquad\left.\begin{array}{r}\text{Targets}\end{array}\right\}\text{Metrics}$$

$$\text{Strategy}\left\{\begin{array}{l}\text{Business Drivers}\\\text{Business Enablement}\\\text{Business Strategy}\\\text{Strategy}\\\text{Strategy \&}\\\quad\text{Program Management}\end{array}\right.\qquad\left.\begin{array}{r}\text{Analysis}\\\text{Analysts}\\\text{Analytical Techniques}\\\text{Analytics}\\\text{People}\end{array}\right\}\text{Skills}$$

$$\text{Agility}\left\{\begin{array}{l}\text{Program Management}\\\text{Information Management}\end{array}\right.\qquad\left.\begin{array}{r}\text{Architecture}\\\text{Infrastructure}\\\text{Technology}\\\text{Tools}\end{array}\right\}\text{Technology}$$

Figure 4.1: Identified model dimensions

This brings us to the following 8 dimensions:

- **Data**
  The fuel for all data-driven activities. How do you source and manage your data?
- **Metrics**
  The key to measuring output and managing performance. How do you use, collect, and enrich your KPI's?
- **Skills**
  Essential for operating a data-driven organization. Do you hire and educate the right people?
- **Technology**
  The foundation for a data-driven organization. What technology do you need to build an analytical process?
- **Leadership**
  The cornerstone for a successful analytical transformation. How does leadership successfully steer the transformation?
- **Culture**
  The driving force behind a data-driven organization. How does culture affect and promote data-driven adoption?
- **Strategy**
  The plan for success. What role does analytics have in your plans and vision of the future?
- **Agility**
  The ability to adapt and deliver. How well are your roles and processes organized to change

and deliver?

4.2.1 Beyond Analytics

One of the goals of this thesis is to place analytics in the context of the whole organization. The maturity models we have seen are focused on analytics as an isolated activity within the organization. The relation between analytics and the organization is often left out or implicitly included in some dimensions. We want to make this relation explicit by defining two more dimensions: *Integration* & *Empowerment*. These dimensions are aimed to position analytics as an activity within the organization as a whole. *Integration* is the concept of integrating analytically produced insight into business processes. *Empowerment* is the concept of empowering members and products of the organization using analytical techniques and data. Data-driven organizations do not isolate analytics in a certain process or department. The entire organization should be better equipped for success through the power of data and analytics.

- **Integration**
  The integration of analytical insight into processes. How is the organization using and integrating analytical output?
- **Empowerment**
  The empowerment of the organization. How is data analytics helping your employees and products to succeed?

4.3 Composing Stages

The different maturity models we have seen have many different stages. They are almost always using unique terminology. Some are focused on communicating business value (e.g. *Business Optimization*, *Data Monetization*), while others focus on analytical capabilities (e.g. *Collect*, *Discover*, *Advise*). Some are describing the organization (e.g. *Analytically Impaired*, *Analytical Competitors*), while others describe what the organization does (e.g. *Discover*, *Advise*).

Some general trends can be observed. All stages start out small or unaware with both a low level of sophistication and a low level of support and adoption (*Analytically Impaired*, *Ad Hoc*, *Unaware*). The last stage is often a stage wherein the organization is in full support of analytics and uses it as a competitive advantage (*Breakaway*, *Analytical Competitors*). The general trend follows the sophistication of analytical capabilities: from collecting and reporting data, to using data for advising future moves.

For the composition of our stages, we want to convey both business value as well as analytical capability. We also want to convey what sets our data-driven maturity model apart from an analytics maturity model. We developed the following five stages with these criteria in mind:

- **Reporting**
  Visualize existing data and create the foundation for an analytical future.
- **Analyzing**
  Dive deeper into the data to achieve insight into why things happened.
- **Optimizing**
  Optimize business processes by integrating analytical insight.

- **Empowering**
  Empower employees and products with data and analytical capabilities.
- **Innovating**
  Use data and experiments to innovate in products and transform the organization.

These stages fit all our criteria. They convey both business value and analytical capability. The stages also convey a holistic view of data-driven activities within the organization by stressing the importance of *optimizing* processes with analytical insight and *empowering* members of the organization. These two stages culminate in an *innovating* organization based on data and analytics.

Chapter 5

Maturity Assessment

5.1 Purpose

To construct an assessment and meta-evaluation, we first have to identify the purpose of the assessment. This vision will help us justify decisions made in the design process.

The goal of our overall maturity framework is to help organizations get started with data-driven maturity. We have noticed that organizations often lack the starting points for a data-driven transformation, and lack the knowledge to bring theory into action. The goal of our maturity assessment is to capture the current state of the organization. When we formulate change, we start with the as-is situation and formulate a plan to reach the to-be situation. The as-is situation provides a convenient starting point for formulating the necessary changes.

Some priorities have to be formulated. We value action over comprehensiveness. While comprehensiveness is important, we believe it is more important to deliver a framework that drives action instead of a very fine-detailed assessment that takes into account a lot of nuances but fails to provide any actionable insight. We value accessibility over accuracy. The assessment should be accessible: easily administered, easy to comprehend, and does not take a lot of valuable resources such as time or money. Accuracy is of importance too, but our aim is to accommodate as many organizations as possible and drive them to action. We can never achieve perfect accuracy, so we aim for Pareto's principle: achieving 80% of the accuracy with 20% of the work.

To define the as-is situation of the organization, we have to assess the current level of data-driven maturity. To make an assessment, we need to evaluate every dimension and determine the current maturity level. Our second goal is to validate our model and our assessment. We do this by conducting an additional questionnaire about the assessment and the results. We will refer to this secondary assessment as the meta-evaluation.

5.2 Design

In this section, we will discuss how we designed our data-driven maturity assessment. We discuss the design decisions made and general concerns for assessments that have to be taken into account.

5.2.1 Quantitative versus Qualitative

We can create an assessment in many different ways. We can opt for a quantitative approach or a qualitative approach. With a quantitative approach, we try in a predetermined and structured way to score the different dimensions using a questionnaire and assessment model. In this way, the organization can also perform the assessment without the assistance of an expert on the model. A qualitative approach consists of determining the level of every dimension with interviews or by self-assessing. This can either be done by a model expert conducting interviews with members of the organization, or the organization self-assessing its maturity level.

Quantitative:

+ Requires no intricate knowledge of the maturity model by the assessor.
+ Assessment fast to conduct.
+ Easily quantified; the result can be in decimals.
− Requires more work to create the assessment model.
− Possibly less accurate, unable to take into account any nuances.

Qualitative:

+ Accurate; takes into account nuances.
+ No questionnaire needed.
− The assessment takes more time.
− Requires maturity model expert.
− Harder to quantify; result usually more general.

A hybrid approach can be done by taking a quantitative result and filtering out further nuances with interviews. However, this still takes a long time and requires interviews conducted by a maturity model expert. Our goal is to provide a quick starting point for an organization and not to compare different organizations. A quantitative approach provides a lower entry barrier because no time-consuming assessment has to be taken. Accuracy is of lower importance because we are not comparing organizations. Therefore, we opted for a quantitative assessment by questionnaire.

5.2.2 Question Types

We use multiple-choice questions to determine the level of maturity. Multiple-choice questions are fast to answer and easy to measure and quantify. Every dimension will feature multiple questions covering the emerging theories of the dimension. The possible answers are scaled according to their corresponding maturity level. All questions are in the form of a list of multiple statements that progresses from the reporting stage to a later stage. Not all questions progress to the Innovating level, because they have 'fully matured' at an earlier stage. This means that with multiple questions per dimension some will progress to the Innovating level and some will stop earlier.

5.2.3 Collection Method

Questionnaires will be administered digitally through a web application. Digital collection has the advantage of being easy to distribute and to participate. The collection and processing of data are easier. At the end of the questionnaire, we can immediately show the results of the maturity assessment in a generated report. We have built a custom web application to facilitate this functionality. A custom web application also allows us to accommodate the questions for the meta-evaluation in a convenient way.

5.2.4 Scoring Model

We will score dimensions on a range from 0 to 5 which indicates the maturity stage of the dimension. A 0 indicates starting maturity and 5 indicates maturity on an Innovating level. The formula for the score takes into account that not all questions progress to the highest level.

$$total(D) = \sum_{q \in D} score(q) \tag{5.1}$$

Where D is the set of questions for a particular dimension

$$score(q) = \sum_{i=1}^{q_a} \frac{1}{D_i} \tag{5.2}$$

Where:
q_a is the answer to question q with possible values [1..5]
D_i is the number of questions in the dimension that progress to stage i

For example, dimension D with 2 questions A and B. A progresses to stage 3 and B to stage 5. The answer for A is 2 and for B is 4. The total score for D is the sum of scores for A, B. The score for A is $\sum_{i=1}^{2} \frac{1}{D_i} = \frac{1}{2} + \frac{1}{2} = \frac{2}{2}$ and the score for B is $\sum_{i=1}^{4} \frac{1}{D_i} = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{1} = \frac{5}{2}$. The total score for D is therefore $\frac{7}{2} = 3.5$

5.2.5 Evaluation Standards

As discussed in Section 2.4, we will use established evaluation standards for the creation of our assessment and meta-evaluation. These standards offer us best practices and help us to avoid common pitfalls in assessment design. In this subsection, we discuss how we have applied these standards in our assessment design.

The discussed standard around found in the paper *Meta-assessment: Evaluating Assessment Activities* [21] by John C. Ory. In this paper Ory discusses student general evaluation standards set by the Joint Committee on Standards for Educational Evaluation [22].

Utility: Audience Identification

"Audiences involved in or affected by the assessment should be identified, so that their needs can be addressed."

Our audience, the parties interested in the result of the assessment, consists of members of the assessed organization. One of our hypotheses is that they seek more information on what it entails to be data-driven and how they can become more data-driven. We address these needs by providing information about the dimensions of the data-driven maturity model, which shows the scope of being data-driven, and the stages of the model, which show more insight in how to become data-driven. We compile a custom report based on the individual answers given by the respondent, with information relevant to the needs in their situation.

Utility: Evaluator Credibility

"The people conducting the assessment should be trustworthy and competent to perform the assessment so that their findings achieve maximum credibility and acceptance."

As discussed before the assessment will be administered electronically without the use of an evaluator external to the organization. The organization is self-assessing, so they are responsible to find an internal evaluator that is trustworthy and competent. The ideal evaluator should have the best overview of the current state of data-driven maturity. In most organizations, this will likely be someone in the role of Chief Information Officer or Chief Technology Officer. The organization also has the option to administer the questionnaire to multiple people in different roles to validate the outcome.

Utility: Information Scope and Selection

"Information collected should be of such scope and should be selected in such ways as to address pertinent questions about the object of the assessment and should be responsive to the needs and interests of specified audiences."

All information that is collected pertains to the data-driven maturity of the organization, with the purpose of evaluating maturity. In our meta-evaluation, we will ask how relevant and important evaluators think questions are, in order to validate this standard.

Utility: Valuational Interpretation

"The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear."

The purpose of the assessment is to form an introspective insight into the current data-driven maturity of an organization with the goal of becoming more mature. The goal is not to compare maturity between organizations or to use the results of the assessment as a performance indicator. Doing so may result in an optimistic bias or a dishonest evaluation. Our assessment offers no incentive for a specific maturity level, as to encourage honest evaluation.

Utility: Report Clarity

"The assessment report should describe the object being assessed and its context, and the purposes, procedures, and findings of the assessment so that the audiences will readily understand what was done, why it was done, what information was obtained, what conclusions were drawn, and what recommendations were made."

We will use these criteria in formulating our report, we might not fully include the technical procedures such as the calculation of the score, as they are of no direct relevance to the key audience.

Utility: Report Dissemination

"Assessment findings should be disseminated to clients and other right-to-know audiences, so that they can assess and use the findings."

Immediately after finishing the assessment the report is available to the evaluator, who has the responsibility to share this with the rest of the organization. The report is evaluated on its usefulness in the meta-evaluation.

Utility: Report Timeliness

"Release of reports should be timely, so that audiences can best use the reported information."

The assessment report is available immediately after the questionnaire has been filled out. The scores are automatically calculated and the report is automatically generated. Quick feedback is one of the advantages of a quantitative approach combined with digitally administering the questionnaire.

Utility: Evaluation Impact

"Evaluations should be planned and conducted in ways that encourage follow-through by members of the audiences"

Follow-through is one of the most important concerns for the participating organizations. We want to make this as accessible as possible by providing a report with information tailored to their maturity situation with the most relevant information.

Feasibility: Practical Procedures

"The assessment procedures should be practical, so that disruption is kept to a minimum, and that needed information can be obtained."

Digitally administering the assessment with multiple-choice questions keeps disruption to a minimum. No extensive interview processes are necessary and no expertise on the model is required. The aim is to keep the required time for the assessment under 30 minutes.

Feasibility: Political Viability

"The assessment should be planned and conducted with anticipation of different positions of various interest groups so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail assessment operations or to bias or misapply the results can be averted or counteracted."

We anticipate the interest of those in organizations who wish to become more data-driven, and we attempt to accommodate their needs by providing relevant information. We will clearly state the use of the results of the assessment: to become more data-driven. We warn to blindly use the results for comparison to other organizations or to use the results as performance indicators.

Feasibility: Cost Effectiveness

"The assessment should produce information of sufficient value to justify the resources extend."

We ensure this by increasing the information value and reducing the necessary resources. We increase the information value in the report by using the assessment results to find the most relevant information for that particular situation in the full maturity model. We significantly reduce the necessary resources by conducting a multiple-choice questionnaire that takes under 30 minutes instead of lengthy interviews with a model expert.

Propriety: Formal Obligation

"Obligations of the formal parties to an assessment (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or to formally renegotiated."

We deem that a formal negotiation is not necessary because the assessment is conducted by an internal evaluator to the organization. A quantitative assessment with fixed questions offers no possibility to stretch the boundaries of the assessment to something outside the intended scope.

Propriety: Conflict of Interest

"Conflict of interest, frequently unavoidable, should be dealt with openly and honestly, so that it does not compromise the assessment processes and results."

Our interest is in that of providing an objective way to assess data-driven maturity, and not to potentially distort results in order to push additional services. The evaluator might have a conflict of interest within the organization; either to overvalue the efforts made by them for the organization or to undervalue the efforts to, for example, create a sense of urgency. We will address this in the report in a section on how to interpret the results, so all audiences are aware of the potential biases of the evaluator.

Propriety: Full and Frank Disclosure

"Oral and written assessment reports should be open, direct, and honest in their disclosure of pertinent findings, including the limitations of the assessment."

All results of the assessment are to be found in the report, with a section on how to interpret the results, including the limitations of the assessment.

Propriety: Public's Right to Know

"The formal parties to an assessment should respect and assure the public's right to know, within the limits of other related principles and statutes, such as those dealing with public safety and the right to privacy"

This standard does not apply to our situation. We will not publicly share the results of the assessment. How to share the results is up to the assessed organization.

Propriety: Rights of Human Subjects

"Assessments should be designed and conducted so that the rights and welfare of the human subjects are respected and protected."

We do not use human subjects for our assessment. We promise the anonymity and confidentiality of assessment results.

Propriety: Human Interactions

"Assessment personnel should respect human dignity and worth in their interactions with other people associated with an assessment."

We will respect the organization being assessed, and are careful with potentially damaging information.

Propriety: Balanced Reporting

"The assessment should be complete and fair in its presentation of strengths and weaknesses of the object under investigation so that strengths can be built on and problem areas can be addressed."

The results offer full disclosure of the strengths and weaknesses with relevant information on how to go further from there. Different perspectives may still exist, the assessment and scoring are based on our own perspective. We do not do a qualitative assessment where we can take into account many different situational nuances. We will try to validate our perspective in the meta-evaluation, as well as warning audiences about possible nuances.

Propriety: Fiscal Responsibility

"Assessment allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible."

This standard is not applicable to our assessment.

Accuracy: Object Identification

"The object of the assessment (program, project, activity) should be sufficiently examined so that the form(s) of the object being considered in the assessment can be clearly identified."

Determining the scope of data-driven activities is one of the focuses of this thesis. We provide a description for each component of the object being studied (data-driven activities within an organization). We will validate if the boundaries of the studied object are clear for the evaluator.

Accuracy: Context Analysis

"The context in which the program, project, or material exists should be examined in enough detail so that its likely influences on the object can be identified."

We aim to provide a holistic view of data-driven activities within an organization. We cannot beforehand determine contextual factors for every organization assessment. Taking organizational context into account when interpreting results is up to the audience, who have the best knowledge of the context.

Accuracy: Described Purposes and Procedures

"The purposes and procedures of the assessment should be monitored and described in enough detail so that they can be identified and evaluated"

Purpose of the assessment should be clearly stated both in the introduction and in the final report. The potential use of the gained information is stated to define the purpose of the assessment and to attract volunteers for conducting the assessment.

Accuracy: Defensible Information Sources

"The sources of information should be described in enough detail so that the adequacy of the information can be assessed."

The source of information will be the evaluator doing the assessment. We will include their name in the report as the evaluator and source of the information.

Accuracy: Valid Measurement & Reliable Measurement

"The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the interpretation arrived at is valid for the given use"

We make use of a standard multiple-choice questionnaire. In our meta-evaluation, we validate the used questions.

Accuracy: Systematic Data Control

"The data collected, processed, and reported in an assessment should be reviewed and corrected so that the results of the assessment will not be flawed."

Data collection, processing, and reporting are automated in custom made software. We have tested this software for bugs to eliminate flaws in data collection and reporting.

Accuracy: Analysis of Quantitative Information

"Quantitative information in an assessment should be appropriately and systematically analyzed to ensure supportable interpretations."

We reduce our assessment results to an overall maturity score and maturity scores for every dimension. These results are visually represented in charts for easy interpretation. This score derivation is explained in the scoring model section.

Accuracy: Analysis of Qualitative Information

"Qualitative information in an assessment should be appropriately and systematically analyzed to ensure supportable interpretations"

We do not collect qualitative information in the main assessment.

Accuracy: Justified Conclusions

"The conclusions reached in an assessment should be explicitly justified so that the audience can assess them."

Conclusions in the assessment are justified by explicitly showing the score for every question and dimension.

Accuracy: Objective Reporting

"The assessment procedures should provide safeguards to protect the assessment findings and reports against distortion by the personal feelings and biases of any party to the assessment."

Nothing can stop the organization to distort the findings in the report. We can warn that doing so is not in their best interest.

5.2.6 Questionnaire

From the information in the full maturity model we constructed maturity questions that determine the maturity level for every dimension. For every dimension, we derived two or three attributes that show a clear progression from no maturity to full maturity. These attributes reflect the full scope of the dimension. We then derived maturity statements corresponding to every model stage for each attribute. The first statement corresponds to a state that precedes the Analyzing stage, and corresponds to 'zero maturity'. The last statement corresponds to full maturity.

Data

Analytics data storage

0 Analytical data is not saved.

1 Analytical data is mainly kept in spreadsheets on local computers.

2 Analytics data is stored in siloed databases across the organization.

3 Analytics data is stored in a central data warehouse with an extract, transform, and load (ETL) process.

4 Analytics data is stored largely in its original format in a data lake, where it can be utilized by different analytics applications across the organization.

5 Analytics data is stored in data lakes but smart IoT devices also collect, process, store, and communicate data.

Source variety

0 Data is not sourced for analytical purposes.

1 Data is sourced from existing operational systems.

2 In addition to sourcing data from existing systems, data with the purpose of analytics is collected where possible.

3 Large scale data about the behavior of users is being collected for analytical purposes and optimization.

4 Third-party data is bought to build a better customer profile. Data is being collected to empower employees in their daily work and to continually improve.

5 Unstructured data, such as voice and images, are analyzed on a massive scale. Innovative products are brought to market with new data collection capabilities.

Governance

0 No data is collected yet, no governance needed.

1 Data is collected on a small scale, without any governance in place.

2 Data is collected and ad hoc governance structures are in place.

3 Basic data governance is in place to help ensure data quality and security.

4 Advanced data governance is in place to ensure quality, security, and correct access over the complete life-cycle of the data.

5 Data governance ensures correct data management over the complete data life-cycle, not inhibiting but accelerating data-driven innovation.

Metrics

Sophistication

0 Metrics and KPI's have not been formulated and put to use.

1 Metrics report plain performance measurements.

2 Metrics aggregate measurements into trends.

3 Metrics combine many performance measurements to provide a better overview of performance.

4 Metrics are backed by analytical models. Analytical metrics are more useful to business users and are better suited for steering the organization.

5 Metrics are analytical and not only used to track existing business but are used in new ways to facilitate and manage new ventures.

Timeliness

0 Metrics are not available.

1 Metrics are only available after the fact and on scheduled times.
2 Metrics are available when necessary and on-demand.
3 Metrics are available in real-time.
4 Metrics are available in real-time and are tracked to provide insight into organizational performance.
5 Metrics are tracked and analyzed to predict future performance when considering multiple influencing factors.

Usage

0 Metrics are not in use.
1 Metrics are used for reporting purposes.
2 Metrics are used for generating insight in past performance.
3 Metrics are used for real-time insight that allows a short improvement and optimization cycle.
4 Metrics are extensively used to empower employees in their daily work and to improve their work and processes over time.
5 Metrics are used to rapidly innovate and start new successful ventures.

Skills

Analytical skills

0 We need the skills to plan and manage a data-driven program.
1 We need information analysis and data visualization skills to create reports and dashboards.
2 We need basic data mining and analytics skills to gain insight about the past.
3 We need machine learning skills to build analytical models for optimizing business processes.
4 We need high-level computer science and artificial intelligence skills to create analytical products and empower employees.
5 We need knowledge on how to rapidly create new analytical models for many different applications and ventures in order to capture the maximum value out of our data.

Focus

0 We have not started any analytical efforts.
1 Our focus is on describing; what happened?
2 Our focus is on diagnosing; why did it happen?
3 Our focus is on predicting; what will happen?
4 Our focus is on prescribing; how will we make it happen?
5 Our focus is on innovating; what can we make happen?

General skills

0 Employees have no specific skills that they can use to get more out of data.
1 All employees in the organization have very basic data literacy skills and can identify what data is relevant.

2 All employees in the organization have basic data literacy skills that make them more comfortable handling data and interpreting statistics.

3 All employees in the organization have statistical knowledge they use to set up experiments.

4 All employees in the organization have the knowledge to continuously improve their work and processes, using a variety of data sources and methods.

5 All employees in the organization have the knowledge to innovate in new products and services using data and advanced analytics.

Technology

Infrastructure

0 No existing infrastructure exists to support analytical activities.

1 Spreadsheets and ad hoc scripts are used to support analytical activities.

2 ad hoc systems and infrastructure exist to support analytical activities and processes.

3 A basic analytical platform is in place to support analytical capabilities and is able to provide production systems of analytical insight.

4 An easily accessible and scalable analytical infrastructure exists, supporting many analytical capabilities.

5 A modular, scalable, and easily accessible infrastructure is in place to support the rapid development of new features and services.

Capabilities

0 We have no dedicated analytical systems.

1 We are able to report data.

2 We are able to analyze data and report the findings in easily viewed dashboards.

3 We are able to optimize business processes and integrate analytical insight in our production systems.

4 We are able to rapidly integrate analytical insight into our business processes and to personally tailor customer interactions.

5 We are able to rapidly build new analytical models and integrate analytical insight into new products and services.

Leadership

Activities

0 Leadership is learning more about the capabilities of a data-driven organization.

1 Leadership is starting analytical activities and making sure the necessary resources are available.

2 Leadership is encouraging analytical initiatives and creating a broader support.

3 Leadership is leading by example by adopting a data-driven process on the highest level and encouraging a data-driven way of working.

4 Leadership is empowering everyone in the organization to engage in a data-driven way of working and to continuously improve.

5 Leadership is embracing new innovation made possible by analytics and lead the transformation to an analytical organization.

Attitude

0 Leadership has little knowledge of analytics.

1 Leadership has an interested attitude toward analytics.

2 Leadership has an exploring attitude, looking for new opportunities for analytical projects.

3 Leadership is convinced of the merits of analytics and embraces a data-driven way of working.

4 Leadership is committed to empowering employees with data and bringing a data-driven way of working to the whole organization.

5 Leadership is passionate about analytics and this is reflecting in every aspect of the organization, and its products and services. Leadership is striving to create a data-driven corporate identity.

Culture

Attitude

0 Employees are unaware of the possibilities of analytics.

1 Employees are skeptical of the benefits of a data-driven approach.

2 Some employees are interested in using analytics while others are hesitant to diverge from their classic way of working.

3 Most employees are using and seeing the benefits of a data-driven approach.

4 Employees are engaged in getting the most out of the data and analytics.

5 Employees are actively promoting analytics to other employees, and culture is the most important driving force for a data-driven way of working.

Adoption

0 The unawareness of employees results in no initiative being taken.

1 The skepticism of some employees can be an inhibitor to data-driven adoption.

2 While some employees start experimenting, others struggle with the adoption of evidence-based decision-making over the old intuition-based methods.

3 Many employees use data and analytics to continuously improve their work and processes.

4 Employees feel empowered by the new analytical culture in the organization, and are seeking new ways to exploit the newfound capabilities in their work and processes.

5 The analytical culture within the organization is a true source of sustainable competitive advantage, employees are always seeking new opportunities.

Strategy

Role of analytics

0 Analytics is not featured in the strategy of the organization.

1 Analytics has an ad hoc and experimental place in the strategy of the organization.

2 Analytics and data-driven adoption have an official place in the strategy of the organization.

3 Analytics has proven to be of great importance to the organization, and is major cornerstone of the organization-wide strategy.

4 Empowering all employees to get the most benefit out of data and analytics is the major focus of the strategy.

5 Analytics is the major driving force of the organization, and the unlocked capabilities provide a competitive advantage.

Strategic focus

0 No strategy is formulated with regard to analytics and data.

1 With a lack of official strategy the focus is on ad hoc analytics initiatives and reporting processes.

2 The strategic focus is on initiatives to formalize analytical processes, and to use data to analyze past performance.

3 The strategic focus is to use analytics and machine learning to radically improve processes and services.

4 The strategic focus is on empowering employees, and to directly integrate analytical insight in the product offerings and business processes.

5 The strategic focus is to use analytical processes to rapidly innovate and possibly transform the business and its offerings.

Agility

Process maturity

0 Analytical processes are non-existing.

1 Analytical processes are largely informal and ad hoc in nature.

2 Analytical processes are formalized.

3 Analytical processes are formalized, standardized, and optimized.

4 Best practices and standards are shared across the organization, processes are capable of continuously improving and restructuring.

5 Analytical processes are capable of facilitating and managing innovation, ensuring a high success rate and little waste.

Roles and responsibilities

0 There are no roles and responsibilities.

1 Roles and responsibilities are not clear and inhibit growth and adoption.

2 Roles and responsibilities are defined more clearly so change can be better managed.

3 Roles and responsibilities are subject to optimization and restructuring efforts.

4 Roles and responsibilities are clear and streamlined, and help everyone to get the most out of data and analytics, while still being flexible and responsive.

5 Roles and responsibilities are continuously evolving to accommodate new innovative business processes.

Integration

Method of integration

0 There are no analytical results to integrate.

1 Analytical results are found in reports that are manually relayed to decision-makers.

2 Analytical results are integrated in business processes with automated reports and dashboards.

3 Analytical insight is integrated in production systems, providing data-driven decision capabilities to products and services.

4 Analytics is centrally coordinated, but locally distributed and available, the integration of analytical capabilities for new opportunities can be done without a lot of effort.

5 Analytics is integrated in all new ventures, new smart products have analytical capabilities integrated.

Action capabilities

0 There are no analytical results to take action on.

1 Decision-makers viewing reports control the performance of the organization and propose new changes.

2 Deeper analysis of the available data brings new problems to light that can be eliminated.

3 Analytical systems predict what will happen, so we can take action to pursue or avoid.

4 Analytical systems advice us what actions will maximize our results.

5 Analytics tells us what to do on a macro-scale; i.e. what business ventures we should pursue, and on a micro-scale; i.e. what product can we best recommend to this particular customer.

Empowerment

Tools

0 Employees across the organization do not have access to any analytical tools.

1 Employees across the organization have access to reports with basic statistics and performance information.

2 Employees across the organization have access to reports and dashboards with insight into the performance.

3 Employees across the organization have tools to set up and manage experiments to optimize products and processes.

4 Employees across the organization have access to self-service business intelligence capabilities, and can use basic analytics capabilities.

5 Employees across the organization have access to advanced autonomous analytics services that help to automatically analyze data and build analytical models.

Education

0 Employees across the organization have no analytical knowledge.

1 Employees across the organization have little analytical knowledge and are learning what information is necessary and relevant to their work.

2 Employees across the organization are becoming more data literate; they can understand reports and statistics.

3 Employees across the organization know how to scientifically approach problems, set up experiments, and validate hypotheses.

4 Employees across the organization know how to do basic data analytics, and how to continually improve the processes under their responsibility through waste reduction and analytics.

5 Employees across the organization know how to use data-driven capabilities in their products and processes, they know how to capture the maximum value out of the available data, and how to source new relevant data.

Information

0 Information is unavailable most of the time.

1 Information is only available through an ad hoc request for information.

2 Information is available in predetermined reports, made available at specific times.

3 Information is compiled on-demand and always available.

4 Information is available in real-time according to own requirements.

5 Information needs of employees are anticipated and relevant information is pushed to employees.

5.3 Meta-Evaluation

To validate the maturity model and maturity assessment, we have to evaluate the assessment practice. We do this by accompanying the maturity assessment with a meta-evaluation. Assessors take this meta-evaluation in parallel to the maturity assessment.

For our meta-evaluation questionnaire, we use two question types. The first type is a Likert scale. A statement is posed for which the scale provides five possible responses:

- Strongly disagree.
- Disagree.
- Neutral.
- Agree.
- Strongly agree.

These responses allow us to measure the sentiment on a certain statement. The Likert scale is also chosen because it is a familiar interface used in many questionnaires. A familiar structure allows the respondent to understand and answer the question faster.

The second question type is a question with options on a rating scale from 0 to 10. Where the Likert scale measures sentiment around a neutral point, negative to positive, we would also like to have the option to measure from 0 to 10 instead of sentiment. The 0 to 10 scale is easier to comprehend because 0 out of 10 indicates zero correlation better than 1 out of 10 would.

5.3.1 General information

We first want to elicit some data about where the organization is currently standing regarding data-driven maturity. We have therefore included the following questions at the beginning of the assessment.

Questions:

- *My organization is planning to become data-driven.* (Likert)
  Does the organization have the intent to become data-driven?
- *We have a clear idea on how to become data-driven.* (Likert)
  Do our model and assessment fulfill a need of the organization?
- *We can use more information on how to become data-driven.* (Likert)
  Can our model help in providing more information?
- *How data-driven is your organization now?* (rating)
  What is the perceived level of data-driven maturity within the organization?

5.3.2 Stages

We first want to evaluate the stages of the maturity model. We will ask if the purpose is clear and if respondents agree with the inclusion of the stage.

Questions (for every stage):

- *The purpose of this stage is clear.* (Likert)
  Have we clearly defined and communicated why this stage is included?
- *How important is this stage in data-driven maturity?* (rating)
  Is this stage relevant and important or can it be left out?

5.3.3 Dimensions

After the questions about the stages, we move to the dimensions. Questions are asked per dimension. In our meta-evaluation we want the sentiment on the purpose, the scope, and if the dimension should be included.

Questions (for every dimension):

- *The purpose of this dimension is clear.* (Likert)
  Have we clearly defined and communicated why this dimension is included?
- *How important is this dimension for data-driven maturity?* (rating)
  Is this dimension relevant and important, or can it be left out?

5.3.4 Attributes

We would like to know, for every dimension attribute, if the respondent has enough knowledge about the organization to answer this question, and if the attribute is relevant for assessing data-driven maturity.

Questions (for every assessment question):

- *I have enough knowledge about the organization to answer this question.* (Likert)
  Can we reasonably assume the assessor is suited to assess the organization?
- *How important is this attribute for determining data-driven maturity?* (rating)
  Are some attributes not relevant or important?

5.3.5 Model, Results & Report

Finally, we would like to validate the understanding of the maturity model, and the usefulness of the generated report. We would like to know if the respondent feels the results are accurate and actionable.

Question (for every dimension in the report):

- *How accurate is the maturity score for this dimension?* (rating)
  Does the assessor feel the result for this dimension is correct?

Questions:

- *The purpose of the Data-Driven Maturity Assessment is clear.* (Likert)
  Is the purpose of the assessment clear to the users?
- *The scope of the Data-Driven Maturity Model and Assessment are clear.* (Likert)
  Have we clearly defined and communicated the scope of the model?
- *The results of the Data-Driven Maturity Assessment can help us create a strategic plan for data-driven maturity.* (Likert)
  Are the results of the assessment actionable?
- *How accurate is the overall maturity score for your data-driven maturity?* (rating)
  Is the overall score accurate?

5.4 Report

The results of an individual assessment are compiled into a standardized report. The goal of this report is to provide an overview of the current state of data-driven maturity, accompanied by relevant information that can form the foundation of a strategic plan. The goal is to keep the report succinct but actionable. We include the following parts in our report:

General overview:

- Overall maturity score and level.
- Name of the organization.
- Name of the assessor; Relevant because the role and experience of the assessor have an influence on the result. A different assessor may find a different result.
- Date; Assessment in the future may yield different results.
- Visual overall maturity score on the maturity chart.

Dimension overview:

- Maturity score per dimension.
- Radial chart with dimension maturity scores; Viewers can immediately see the strengths and weaknesses.

Personalized maturity model:

- For every dimension, we show the relevant part of the maturity model at that maturity stage. For example, the maturity score of the Data dimension is 1.2, indicating a score between Reporting and Analyzing. We then show for the Data dimension the part in the maturity model at the Analyzing level, so the information is the most relevant.
- For every dimension a list of suggested action items for that level of maturity. Using the results of the assessment we retrieve the most relevant action items for the current level of maturity.

An example report can be found in Appendix 11.1.

Chapter 6

Iterative Development

In this chapter, we will dive deeper into the iterative development of our data-driven maturity model. We first discuss the development strategy in Section 6.1. In the remaining sections, we discuss the phases of development, with the adjustments made in each phase.

6.1 Strategy

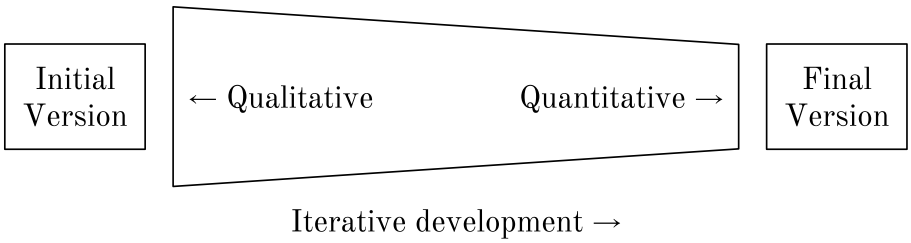

Figure 6.1: Iterative development strategy.

Our approach to iterative development is shown in Figure 6.1. We want to refine the initial version of the model and assessment to the final version. The model goes through a refinement funnel where we incrementally refine the model. In the beginning, we focus more on getting qualitative feedback to both improve the user experience of the assessment and model. Later on, we focus on validating the usefulness of the model through quantitative feedback.

We first distribute the assessment to a targeted audience that is more likely to provide qualitative feedback. This way we can immediately work out the biggest shortcomings before it is available to a broader audience. The broader audience will have access to a refined model. We will use the quantitative feedback of the broader audience to validate the model and assessment.

Our notion of iterative development is the continuous delivery of new changes. This means we continuously push small changes live. We will be able to collect feedback faster. We have grouped all the small changes in phases.

6.2 First phase

The initial version of the model was distributed to a carefully selected audience ($n = 5$). The focus is on gathering qualitative feedback, with a focus on the usability of the assessment. We will discuss adjustments we have made in the model and assessment.

- **Better division between meta assessment and assessment**
  We have received some questions about how the meta-evaluation questions are relevant to the data-driven situation of an organization. Some were under the impression these are relevant for the final score. However, the meta questions are not relevant for the final score, only for gathering feedback. We have added extra information on this in the assessment introduction, and we have added a completely different color scheme for meta questions to highlight the separation between meta-evaluation and assessment.

- **Support for Internet Explorer**
  The assessment application was made with modern web development techniques. Support for Internet Explorer, an older browser with limited support for the latest web techniques, was not a priority. We received feedback about the application not working on IE. IE's market share is 12.4% as of April 2018 according to NetMarketShare [38]. It is still used in many corporate environments, which we do not want to exclude. We made adjustments to be able to run the application in IE.

- **Live support**
  Before and during the assessment, participants had (technical) questions about the assessment. We do not want participants to be stuck, so we added a live support chat with Drift [39], for support or feedback. This way participants would not get stuck during the assessment and we do not lose valuable feedback.

- **Sharing results**
  The final report of the assessment is displayed on a webpage. Participants wanted to save or share this report with colleagues. We have added an option to download the report as a pdf file and to share or access the report with a permanent link.

- **Action list**
  A common frustration we saw, not only in our assessment but in general, is the lack of concrete advice. Often a lot of theory is spouted without any advice on how to apply it to your organization. Our model already featured specific advice, relevant for your current maturity, but we have also added a bullet-point list with actionable items.

- **Definitions**
  Another point of feedback was that some terminology was unknown to the user. We have added definitions of a list of concepts when the user hovers the concept. This way users do not have to leave the application to find a definition, and we can more accurately specify how we interpret a certain concept.

- **Textual refinements**
  We have received small textual refinements such as typos. We also received feedback on some sentences which were not very clear and now have been adjusted.

- **Execution to Agility**
  We have decided to change the Execution dimension to the Agility dimension to better reflect the purpose of the dimension. Both names refer to bringing strategy to life, but Agility better encompasses the change management necessary to make this a reality.

- **Renaming of questions**
  We have renamed 'Leadership - tasks' to 'Leadership - activities' to better reflect the role of leadership during this stage. We have also renamed 'Culture - effects' to 'Culture - adoption'.

6.3 Second Phase

The second phase is distributed to a small group of experts ($n = 4$) on analytics and data maturity who can provide us with additional qualitative feedback, with a focus on the content of the model and assessment. In addition to smaller content refinements, the following changes were made:

- **Data volume**
  After an initial analysis of the survey responses, the most polarizing dimension is the Data dimension. The standard deviation on the question 'How important is this dimension for data-driven maturity?' is the highest of all dimensions (sample standard deviation $s = 2.82$). The sample mean of the same question is 6.75. The most controversial attribute of the data dimension is the volume attribute, $s = 3.28$ & $\bar{x} = 5.25$ for the sample standard deviation and the sample mean respectively for the question about the importance of the data volume attribute.

  The qualitative feedback also suggested the relative unimportance of data volume. Experts felt that volume should not define how data-driven an organization is, as an organization that is capable of utilizing data close to the source in automated processes is also data-driven. This raises the question of how relevant big data is for being a data-driven organization. After analyzing the quantitative and qualitative feedback, we have removed the volume attribute and questions from the assessment. Large volumes can be considered a complementing data strategy, but should not be a goal. Automating the process from data to action is held in higher regard than amassing and manually analyzing large data sets.

- **Emphasis on integrating data processes**
  Some experts felt a strong emphasis on business intelligence. While it is true the first two stages have a bigger emphasis on analytics, our model features an Optimization stage and an Integration dimension, both with a large focus on automating analytical insight into business processes. It may be the experts did not get the same view from filling in the assessment. We have added more emphasis on this in the descriptions of the Optimization stage and Integration dimension.

- **Visual update for maturity questions**
  We have updated the maturity questions visually because some participants were not aware they were questions that also have to be answered. Indeed the meta questions in pink standoff visually opposed to the maturity questions. The maturity questions are now visually isolated with a blue background, in order to more easily recognize them as actual questions.

6.4 Third Phase

The third, and final, phase focused on additional validation of the model through additional quantitative feedback ($n = 8$). The meta-evaluation provides us with quantitative feedback, which we will use to statistically validate the components of our maturity model. After this phase, we can conclude if the model holds ground or should be re-evaluated. Participation in this phase was open to everyone, made public through LinkedIn posts, but we also actively pursued potential respondents.

No more major flaws emerged from the third phase that needed to be addressed. Some textual updates or nuances were added to clarify or emphasize. The full analysis of the results can be found in *Chapter 8: Results & Discussion*.

Chapter 7

Final Maturity Model

7.1 Dimensions

7.1.1 Data

Data is essential to all data-driven activities. Data is information, and information can reveal insight. Data fuels all data-driven processes that refine raw data to valuable insight. This is why data is often compared to oil: left untouched their use is limited, but after collection and refinement it has high potential [40]. The inherent properties of big data: volume, velocity, variety, veracity, and value add new challenges to managing data [41].

In the world of big data and the internet of things, there is an enormous amount of data being generated. The first step is to identify which data has *relevancy* and must be collected [42]. This challenging part is determining beforehand what data will be useful or not. You might want to collect online reviews about your company for sentiment analysis. Do you also collect the reviews about competitors? Depending on the goal of the project you choose what data to collect. You also have to decide which parts of the review you wish to collect. Is the name of the reviewer relevant? Probably not, but the time of placement of the review might hold some insight.

Just like oil, data can be hard to find and extract. Where do you *source* your data from? Data can reside in legacy systems without easy access, or it can be bought from third-party vendors. Which sources can you trust and are easily accessible?

Getting data to your analytics platform can be challenge. You need the right connections and infrastructure to integrate data flows in your systems. More on this in the *Technology* dimension. An important aspect is the *timeliness* of the data, especially for automated data-driven processes. Data relevant a week ago may not be relevant today. Data needs to be at the right place at the right time.

Not all data is fit for analysis. You may need to combine data with other sources or *transform* it to a more suitable form. You need to decide when, and how, you would want to do this. It might be more efficient to prepare the data after extraction, rather than to do it when you start the analysis later on. Data is often transformed beforehand so that it will be easier to perform analyses on it afterward, such as the case in *online analytical processing (OLAP)* [43].

Data is an asset, and all assets need to be managed. To govern all these aspects of data, it is important to establish responsibility for data assets. *Data Governance* concerns itself with the management of data during its life-cycle [44]. Many different governance strategies exist, depending on the size of the organization and the volume of data, some may be better suited than others. Some overlapping concerns are data quality, data access, and data life-cycle management.

Ensuring data *quality* is a challenge that is getting more relevant in the age of big data. More data, from an increasing variety of sources, means the data you collect is more liable to containing errors. If the quality of the data becomes too impure, the decisions you make based on that data will become flawed. Processes should be in place to monitor and improve the quality of the data you use for analytics.

Data *security* & *privacy* are two related concepts that can easily be overlooked by new practitioners. The classic way of thinking is that investments in security and privacy are not directly value-adding and are therefore often postponed until it is too late. But losing data or compromising customer privacy is a great financial risk, and mitigating this risk should be a primary concern when managing data. Strict authorization and access control should be in place. User privacy is also no longer an afterthought. Users value their privacy and organizations should act in a responsible matter in order to not breach the users' trust. New legislation, such as the EU's General Data Protection Regulation, also force organizations to comply with standards regarding privacy and data protection [45].

7.1.2 Metrics

If data is about the input, metrics are about measuring the output. You can measure the specific performance of a process or track more global key performance indicators (KPI's) [46].

Determining which indicators are key sets the priorities of the organization. KPI's should reflect the objectives of the business strategy. If you are are aiming for growth you might track total sales per time period, but if you are trying to optimize the digital advertising budget you can look at the return on advertising spend (ROAS).

Metrics should also be used to monitor the organization. Data-driven organizations enable managers to view real-time metrics, which can be used to steer and correct. Timely feedback is important. A real-time dashboard allows managers to quickly identify trends. Trends can be both negative in which case they can be corrected or opportunities which can be seized.

Metrics form the core of the performance management program. They communicate the priorities of the organization and prevent a focus on trivial tasks. Management should be aware of the side effects of a hard focus on metrics and targets. The amount of sales as the most important measurement of performance might incentivize the usage of aggressive promotions 'to meet the targets'. This behavior can hurt profitability in the long-term. Especially when performance bonuses are tied to flawed KPI's.

Collecting some metrics can be a challenge. Some straightforward financial metrics, such as turnover or profit, are usually easy to measure. Some more advanced metrics, such as customer lifetime value (CLV), the prediction of how much profit a customer is going to generate for the entire future, are very hard to correctly derive, and depend very much on the type of business. Deriving meaningful metrics can be a challenge, but the right metrics can greatly impact company performance.

7.1.3 Skills

The people in the organization and the analytical skills they have form a vital part of the analytical process. Talent needs to be developed, managed, and attracted. Data does not work on its own, it needs to be put to work by people with data science and analytics skills. Many organizations starting analytics efforts do not have the necessary skills, and as a result struggle to get the most value out of their data.

Data science is the set of principles and techniques that guide the extraction of information and knowledge from data [47]. Principles and techniques draw from statistics and computer science such as data mining and machine learning. These fundamental skills have to combined with visualization skills to communicate with other parts of the business. Finally a data scientist needs to be able to "view business problems from a data perspective" [47].

Depending on their background employees can follow training for analytical skills but it is a hard subject to get into without an analytical background or education. Formal education is often required for areas such as big data and machine learning. Therefore it is important to attract and retain people with the right set of skills. Attracting talent is a challenge because the demand for data science skills is rapidly growing, which puts pressure on the demand [48].

Data literacy is a basic understanding of what data means and can do, including interpreting statistics and drawing conclusions from data [49]. Data literacy consists of a set of skills which are beneficial to have for all employees, not only those in a analytics department. Everyone should have the ability to correctly interpret results or dashboards made by the analytics department. Data literacy across the organization can be increased by training and hiring people with analytical backgrounds.

A broad range of specific analytical techniques are available with different levels of sophistication and different applications. Based on the goals of the organization different techniques can be used. Very advanced techniques such as deep learning require very specific knowledge about the subject and may require people with specific backgrounds. The required knowledge can be very specific to the domain and organization.

7.1.4 Technology

If Data is the fuel, Technology is the catalyst for data-driven activities. Technology plays a role in every step of the analytics process. Different technologies are used to generate, store, and process data. We can identify 7 different main areas for technology related to data-driven processes: Collecting, Extracting, Storing, Analyzing, Visualizing, Integrating, and Connecting. There are often many different technologies or products that solve the same problems, albeit in a slightly different way. For every problem you have to decide which technology is going to be used. Do you build it yourself? Is there a product on the market that solves our problem within our budget? There is no silver bullet. Depending on your needs and available resources, both money and skills, you will have to decide what the best solution is.

Data is either not being generated yet or residing somewhere you cannot perform analytical processing. For example you want to better understand user behavior on your website, but you have no tracking implementation yet. You would first need to use the technology to collect data. There are many (free) web analytics solutions, so it does not make a lot of sense to build a solution yourself.

Other times the data is already there, for example customer and order data in a CRM system. You still have to extract data from the system to a platform where you can analyze the data. You have to build or use existing connectors to extract the data. You might also need to transform or cleanse data before it is useful for analysis.

The big data revolution gives us enormous amounts of data at our disposal. Storage capacity is continuing to grow while also getting cheaper. You can store data in many ways. You could build big archives in flat files and store them on disks. But in this scenario it would be very hard to analyze or search in this data. For analytical purposes we want to store the data in a database or data store, which provides easier and faster access than the standard file system. Many different database solutions exist that all have different features and specialties. Picking a database solution is often a trade-off. For analyzing it is important to be able to store large amounts of data and having a high throughput to be able to process it all. For example Apache Hadoop Distributed File System (HDFS) [50] is a distributed data store for storing and processing big data on commodity hardware. An implementation of MapReduce [51] then allows parallel processing of big data sets.

When the data is stored somewhere easily accessible you can start to analyze it. There are many tools and techniques that can be used to analyze data. There are many solutions to explore or mine data for insight. Open source machine learning solutions, such as WEKA [52], can be used to mine for insight and to create analytical models that can be used to make predictions. In other scenarios it might be better to create and apply machine learning algorithms yourself.

Visualizing data is an important component of analytics. Presenting data and metrics to those making the decisions allows them to base the decisions on data, instead of intuition. The easier the data is to understand, the easier it is to make decisions based on the data. Many solutions and services exist to visualize the data by making dashboards, charts, and more. New tools, with simplified user interfaces, allow end-users to create their own dashboards with the available data. These new tools allow the business side to become more self-reliant and flexible. The challenge is to get the right data to the right users at the right time, allowing users to make the best decisions at the right moment.

The automatic integration of insight into your business processes allows you to capture the maximum value of your analytics efforts on a big data scale. Machine learning capabilities, like automatically recommending products, provide a tremendous amount of value. Automation means you can make decisions on a micro-scale. For a customer facing company this can mean personalizing interactions for small segments or individual customers. Integrating is a technical challenge with two major challenges. The first challenge is to connect analytics environments with production environments. Data goes in the analytics environment and insight flows out. This insight has to be available in production environments to be utilized in customer interactions. The other challenge is the timing. Manual insight retrieval, such as in reports, is often about historic data, e.g. last month. Automated insight integration often requires rapid insight creation and integration. For example, when a customer views a product you might want to update their profile and show new recommended products. This has to happen within seconds. There is no time for grouping data and performing analytics in a big batch.

The final area in technology is about bringing all components together. Connecting everything is a challenge, because you are often working with a large variety in systems. There is also no standard solution or software that is capable of this. In the beginning a lot of systems will be connected manually, by extracting and importing data, or by batch processes. For our big data and integrated analytics needs this process will need to be faster with a greater capacity. This makes the data pipeline a central component to the data-driven organization. Instead of looking to connect two systems we might look at a central data flow where systems can plug into. A central nervous system of the organization with different components reacting to events.

7.1.5 Leadership

Leadership is an important factor determining the success of an analytical transformation [7].

The first task is to get analytics on the agenda. In the beginning stages, a strong advocate (champion) is required to set efforts in motion. The leader of a business unit, with a strong passion for data-driven decision-making, can provide the ideal starting point for building analytical capabilities. Leadership is responsible for allocating the necessary resources for building an analytics team.

The next task is to continue encouraging the data-driven way of operating, both to other leaders as to employees. Upper management has the best chance of changing the orientation of the company toward a focus on data [37]. They have an important role in communicating results to other leaders, in order to convince them of the merits of analytics.

Several characteristics make a successful analytical leader. Analytical leaders have a passion for using data and analytics in order to drive business success, which they use to inspire the organization [37]. They are not afraid to act on the outcome of analytical processes. They trust the outcome of a solid data analysis process more than their own intuition. Nothing will inhibit analytical growth more than ignoring the results of analytical processes.

7.1.6 Culture

Culture plays an important role in the acceptance and adoption of a data-driven strategy. Analytical culture stretches beyond the analytics department. Some companies such as Google or Amazon are based on analytics, and therefore have nurtured an analytical culture from the beginning. Management in other companies may be more traditional, and rely on their intuition more. There are two main challenges in creating an analytical culture.

The first challenge is the acceptance of data-driven decision-making. Many companies are used to base decisions on intuition and personal knowledge, with the highest manager having the final say. Analytics provides a solution backed by data, but this solution is worthless if it is dismissed by management. Management can have a number of reasons to ignore the findings. They may not trust the analysts or their methodology. This is often fueled by insufficient knowledge about data science. That is why data literacy is important for everyone in the organization. Leadership is important to set the right example to other levels of management and to promote the acceptance of data-driven decision-making.

The second challenge is to engage and empower all employees to adopt a data-driven way of thinking and operating. Every employee should be encouraged to experiment and measure, in order to optimize their work.

7.1.7 Strategy

Strategy determines what role analytics will play in the organization. Every company is different and has different competitive advantages. For an e-commerce company, selling to thousands of customers everyday, it makes sense to use prescriptive analytics to implement a recommender system, while for another company, selling expensive devices to a couple of businesses per month, it would make no sense. There are many applications for analytics and many of them will not align with your goals and business model.

Leadership sets the strategy. The first thing they have to decide is the role of analytics. What are the competitive strengths of the company? How can a data-driven approach sharpen those strengths? What is the long-term vision for analytics in the organization? If you truly want to compete on analytics you have to choose a distinctive analytical capability on which to compete [37]. This will be different for every company. Some excel through reliable logistics, others predict extremely well what customers are going to prefer. The starting point for your analytics transformation will depend on where you think analytics can make the biggest impact.

Strategy will transform over time and change with each stage. Early strategy focuses on proving the worth of analytics or beginning to build analytical capabilities. This should be reflected in, for example, the hiring strategy, by hiring data analysts. Strategy should also determine how the organizational structure is going to look like. You might want to start by setting up a team inside department relating to your distinctive capability and grow from there. Medium-term strategy should focus on continuing to build analytical capabilities, training and retaining data analysts, and creating lasting value for the organization. In the long-term analytics has to become a core part of the strategy. It will become a capability you can compete on. Products or services may be introduced that have analytical capabilities as a feature.

7.1.8 Agility

Strategy focuses on the big picture, Agility focuses on how to adapt, execute, and bring the strategy to life. Too often a strategic plan fails due to rejecting the proposed change. Analytical transformation requires a massive paradigm shift from intuition to evidence-based reasoning. The plans may be clear and understood, but it is still very easy to slip back into the old ways. New roles and power balances are introduced, which can cause a struggle to adopt the new structures. Agility focuses on implementing the strategy and enabling the program to succeed.

Successful strategic execution is the continuous process of aligning the things you do with the things you should be doing. Many issues can prevent an organization from successfully implementing their strategy. In 'The Secrets to Successful Strategy Execution' [53] Neilson et al. identified two majors factors for successful strategy execution. The first factor is clarifying decision rights. Clearly defining the roles, and establishing who is responsible for making a specific decision. For our data-driven transformation this is important because the decision-making responsibility may shift from a manager to a data scientist. If the responsibilities are not clearly defined the manager may attempt to keep overruling the data. It may be necessary to develop a change management process to effectively transform the way the organization works.

7.1.9 Integration

Integration is a concept that is sometimes overlooked when talking about analytics and data science. Integration in this context is about how the product of analytical efforts is flowing back to the business, and how insight is turned into action. Not every insight can be turned into action by the one doing the analytics. You may need a developer or a manager to implement or utilize the insight.

We have manual integration, for example a report delivered to an executive or an opportunity for improvement communicated to a manager. This is often this is the only form of integration. For a team of analysts looking for insight in collected data, or a web analyst looking for website optimizations, there is no automated system integrating the results. Feedback is communicated through reports or presentations, which spur actions that will improve the business. Good manual integration means a process is in place to retrieve and execute on the results of analytics. For example if a web analyst identifies an opportunity for better conversion there should be a process in place to turn this opportunity into action. They should not have to wait indefinitely for implementation. The faster insight is turned into action, the more value is created. Fantastic reports can be created but if there is nobody to read them, or no resources to take action, the analytical efforts will be wasted.

Organized manual integration becomes information and knowledge management. Making sure the right information flows back to the stakeholders. A lot of useful information and experience is gained over time, but this information is not often documented. When the organization increasingly starts to run experiments the results have to be documented and easily accessible.

Automated integration is the process of automating insight into action. Recommender systems are a good example of automated integration. Customer data is ingested into a recommender system. The system turns the data into insight: products the customer is likely to purchase. This insight flows back to the production systems. The next time a customer views a webpage the recommended products are featured on the page. Automated integration is the result of a successful combination of analytics and implementation. Automating this process has many challenges. Systems have to be made to make insight available to production systems. Timing is a big challenge. It is one thing to collect and process the data to insight and make a report or spreadsheet. It is much more challenging if this all has to happen in real-time. The whole analytics process has to be automated and fast enough to deliver in-time.

7.1.10 Empowerment

Empowerment refers to the ability of all employees to engage in analytics. To put the right tools, techniques, and information in the hands of the workers on the ground so they can discover opportunities and keep progressing. Ensure the flow of information goes where it is needed. One of the goals of a data-driven transformation is to make sure everyone has access to relevant information at the right time.

A simple example of employee empowerment would be a dashboard available to a manager. It simply displays aggregated data but the manager is able to look at this and identify trends or problems. This puts the power of analytics directly into the hands of the users. A strong analytical culture and widespread data literacy is key to making this successful. You can also make the data available to the manager. Data sources and visualization tools are available, so a manager can create his own dashboard. Many tools such as Microsoft Power BI [54] are being developed that allow users to create their own analytics solutions. This is called *self-service business intelligence* [32]. The goal is to make the user more self-reliant. They have the best knowledge about what they would want to have available in a dashboard because they have to use it themselves. Empowering the organization encourages them to engage in data-driven activities.

Experimentation and process analysis are activities that can be encouraged and enabled. Employees can be empowered by learning and using process optimization techniques. Everyone in the organization has a stake and a responsibility in creating a stronger organization for the future. Techniques traditionally used in manufacturing, such as Lean, are increasingly used to improve knowledge work processes [55]. These analytical techniques are can be applied universally, even in the age of information. The principles help foster a culture of innovation and continuous improvement that is fueled by data-driven insight.

7.2 Stages

The design of the stages of the data-driven maturity model is more complicated. Not all stages of the analyzed maturity models can be covered literally. We attempt to reflect all underlying capabilities and arguments into the stages. We give a general description of the likely state the organization is in, followed by a more specific description for the dimensions of the model at the current stage.

7.2.1 Reporting

Reporting is the first stage in our data-driven maturity model. The focus in this stage is on orchestrating efforts to set up analytics. The goal is to make data available for analyzing, and do basic reporting. Analytics and a data-driven way of working have to become official. Until now analytics has been done on an ad hoc basis, with no formal processes in place to guide the activities. Spreadsheets are the main tool used for basic analysis and reporting. The proliferation of spreadsheets leads to confusion, as there no longer is a single source of truth or any responsibility for data and spreadsheets.

But there is a growing interest in the potential of using data and analytics to support the decision-making process. Both by employees as in upper management there is a desire to no longer rely on intuition to make decisions. The first step to attaining insight into how the company is doing is to report past performance. This information can be analyzed by managers in order to make decisions regarding the future.

Data

Data is fundamental to starting data-driven activities. Valuable data is hard to access in many different production systems. The critical task is to determine which data is relevant enough to collect. At this point, the number of data sources is limited. You often retrieve the data from your own systems. It can be challenging to retrieve data that resides in functional silo's. Because you are reporting over historical performance, the timing is not a big challenge. Data is often transported in a daily job to a primitive data warehouse. The quality of the data should be kept in mind but is not a huge concern at this point, because the sources are trusted and the variety of the data limited. There are no real big data challenges that must be faced.

- Determine which data needs to be visible.
- Source and report critical data.

Metrics

Metrics are an essential part of reporting. Metrics provide an easy performance overview. Just like determining which data is relevant, the first task is to determine which metrics should be established. Dashboards with near real-time metrics should be made available for management to quickly identify trends. For operational activities dashboards can also be made to spot operational issues. In this stage, the focus should be on easily accessible metrics, which can later be replaced by more advanced derived metrics backed by analytical models. Overall there is a focus on current performance, not yet on long-term progression.

- Determine critical metrics.
- Create basic dashboards with the most important metrics.

Skills

At this point skills will be unavailable or underdeveloped. The focus is on sourcing talent, either internally or externally. Find people with the initiative to build and expand the analytics program in the organization from the ground up. Attracting people at this stage should be a careful process, because the first data scientists are for a large part responsible for the early success of the program. Early success will convince leadership to invest and scale, while early struggles may lead to termination of the program. Early success also spurs the interest of other employees in analytics, and a cultural shift to a more analytical culture.

- Source talent internally or externally to build an analytics program.
- Focus on early success to gather momentum.

Technology

The first step is to identify the business requirements to eventually design or acquire technology capable of delivering those needs. The selection of technology should be a careful process that analyses the many trade-offs. In the beginning stages, it is important to prove the value of analytics so it might be attractive to use a Software as a Service (SaaS) model in the cloud, where the hosting and powering of a system are taken care of by an external provider. This usually requires lower upfront costs and faster bootstrapping. Many technologies are available this way. You can also take on more burden yourself by using an Infrastructure as a Service (IaaS) model. With IaaS the infrastructure, such as servers, are provided on which you install the applications you need. The trade-off is between speed, flexibility, and costs. SaaS is often the easiest to get started with, but may no longer fit your needs over time and costs more. The technology in this stage focuses on the starting steps of an analytical platform where data is collected and made available in reports and dashboard for management. These dashboards provide an overview of historical and current performance, and leave generating insight to the ones viewing the reports.

- Identify business requirements for a basic analytics platform.
- Decide if you want to build or buy software and infrastructure.

Leadership

Leadership has an important role in the early stages. It is responsible for initiating analytical programs and leading the cultural shift. In the beginning there will be a division in upper management with skeptics and believers in analytics. Someone in leadership should champion the idea of analytics and foster the program.

- Initiate analytical program by putting it on the agenda.
- Champion and foster the analytical program at the highest level.

Culture

The starting culture will largely depend on the type of organization. Companies that rely on intuitive decision-making will largely be indifferent to analytics. They are used to make decisions based on gut instinct, with the highest paid employee having the largest influence. This can also lead to opposition to a culture based on analytics. Employees who were used to being valued for their intuition-based decisions may feel threatened by the shift to evidence-based decision-making. It is necessary to remove worries by guaranteeing they will not be replaced by the transition. An excited or open attitude by the entire workforce increases the chances of success of the program.

- Warm up and excite employees for an analytical culture.
- Remove worries about the shift to evidence-based decision-making.

Strategy

The desire to move from intuition-based decision-making to evidence-based decision-making should be reflected in the Strategy during the early stages. Additionally, a long-term vision for analytics should be developed. How will analytics transform the way business is generated in this company? For some companies, this might seem too far into the future, but this exercise provides both a starting point for analytical activities and a future you can work toward. It makes no sense to start the analytical activities with logistics, if you do not plan for it to be a competitive strength in the future. In the beginning, there are two main approaches. Starting small and proving value before moving bigger, or going all-in on analytics. Starting small is a slower approach, but might be an easier transition for the organization and the employees. This approach is also useful if upper management is not yet fully convinced.

- Identify core competencies of the organization.
- Develop a long-term strategic vision for analytics.

Agility

In the beginning the execution of the chosen strategy will be largely informal and ad hoc. Moving fast and proving value is more important than adhering to strict processes. Embrace internal entrepreneurship in the early stages to maximize the early results and excitement. For a smooth start, it is however important to define basic responsibilities. For every decision, it should be clear who is responsible and accountable. Prevent projects and initiatives getting stuck because nobody is sure who is responsible for making the decisions. Without responsibilities it is hard to drive change. Defining roles and responsibilities also helps ease the shift in power balance. A new source of insight is being created, which may support or replace the previous shot-callers. If responsibilities are not clearly defined it may happen that the newly created insight is not properly used, and keeps being overruled.

- Embrace entrepreneurial spirit to move faster instead of defining strict processes.
- Define basic responsibilities and decision domains to avoid getting stuck and to keep driving change.

Integration

Integration is the beginning stages will largely be manual by delivering reports or ad hoc analysis of a specific business question. It is important to cooperate with managers to get the correct questions they need answers to. Standardized reports can be created which can be reused for monthly or quarterly updates. Software packages can provide automated integration for some specific tasks, but often big gains can already be made by manually relaying back insight to the business. Timing is, in this stage, of less importance than correctness.

- Define which information is needed by decision-makers.
- Standardize reports.

Empowerment

Early on employees will not be empowered by analytics. They do not have dashboards, and do not have the opportunities to do things themselves. There is little focus on Empowerment in the early stages of the process. However those responsible for analytics can create dashboards for employees to use, to follow current performance and make decisions.

- Create basic dashboards and reports to empower employees.

7.2.2 Analyzing

The Analyzing stage is where analytics is getting more formal and organized. The birthplace of insight shifts from the manager reading a report to the analytics department. Many of the prerequisites have been taken care of in the Reporting stage. Management is now able to view predetermined reports and KPI's, and use them to make more informed decisions. The focus in this stage is to look deeper into the data, better known as data mining, to find insight that might not be obvious. Analytical models can be created to predict how events will unfold in the future. Learn about the past in order to predict the future.

Data

The focus in the Reporting stage was on making data insightful for users by making it available in reports and operational dashboards. In the Analyzing stage, you go deeper into the data to find insight not visible from the surface. Where you previously collected data already available, you now might want to look at creating additional sources of data. You will start to collect data that is not only relevant for reports, but also for further analysis. Collecting more data, and from different sources, brings its own challenges. Data quality might become an issue if the sources are too unreliable. Monitor data quality and assign responsibility for certain domains to ensure it never drops below a minimum. Security and privacy should be a great concern and a priority. User data should be anonymized and secured.

- Determine which data is relevant for deeper analysis.
- Collect and analyze relevant data.
- Monitor and ensure data quality.
- Define a strategy for keeping data secure and private.

Metrics

In this stage, basic metrics are collected and displayed in dashboards. Because you are collecting metrics for a longer time you can now also track KPI's over time. This provides additional insight over the progression of the organization. Make sure to fully realize what drives the KPI. KPI's are the outcome of a process. Focusing on a correct process will provide long-term results, while artificially inflating KPI's only serves the short-term. You now also have the capability to create KPI's backed by analytical models. For a customer lifetime value (CLV) analysis you need a lot of historical knowledge. If you can create an analytical model that predicts the customer lifetime value for a new customer, you can estimate how much you can spend to attract this customer. The CLV can also serve as a KPI that can be tracked.

- Define more advanced KPI's.
- Track KPI's over time; focus on long-term results.
- Create analytical models to back analytical KPI's.

Skills

Analytics evolves from showing numbers to calculating numbers, and the required skills evolve with it. Data mining and analytical skills are the most important skills in this phase. The current team has to be trained for more advanced analysis and the creation of analytical models that can be used for predicting future events based on new incoming data. In addition to basic statistics, advanced knowledge of data and data mining techniques is necessary. The retention of current talent is also a priority, because it can take a lot of resources to bring newly hired employees up to speed. A small team in the beginning means that a few individuals will have critical knowledge about the systems. Expanding the team will spread the critical knowledge and help overcome a sudden departure.

- Train or attract data scientist to dive deeper into the data.

- Retain talent with critical knowledge.
- Introduce employees to basic data literacy skills.

Technology

Where in the Reporting stage the infrastructure was largely ad hoc, a professional analytics platform will be built during this stage. The implementation details will largely depend on the organization's requirements. There has to be a standardization of technology, and a modular platform that is capable of easily integrating new features and components. All data will be stored, in a standardized format, in a data warehouse or data mart, and is easily accessible for services. Invest in tools that help you analyze and mine the available data. A lot of technology is being developed to help the data scientist in making the available data more explorable and increasing their efficiency.

- Professionalize the data infrastructure.
- Standardize technology, integrations, and warehousing.
- Acquire technology that helps data scientists analyze data.

Leadership

Leadership continues to be an important driving force for data-driven maturity, as long as not everyone in the organization is convinced. Early success has hopefully convinced the majority of upper management. If leadership praises the improvements through analytics, the rest of the organization will more likely embrace the changes. Basing decisions on data means no longer the most influential employees are taking the decisions, but decisions are made based on the correct use of analysis. Leadership has to follow through on this all the way to top level decisions. You can not force low-level decisions to be based on data while continuing to base high-level strategic decisions on intuition.

- Continue to drive home the importance of a data-driven organization.
- Show the results of data-driven efforts to the organization.
- Lead by example and base high-level strategic decisions on data analysis.

Culture

The Culture inside the organization is still lukewarm to the use of analytics. They may know of some activities that are being started, but it is still isolated from the rest of the organization. The successes may inspire some early adopters to also try more experiments to confirm their beliefs. Some struggle may occur between believers in a data-driven way of working, and those who 'have always worked this way', and continue to do so. Clear direction from leadership and management is needed to resolve this.

- Use early success as inspiration to the rest of the organization.
- Set a clear direction from the top down to move away from the old way of working.

Strategy

After carefully setting the first steps toward a data-driven strategy, it is time to expand the efforts and give analytics a permanent place in the organizational strategy. The focus expands from reporting on past events, to also analyze why things occurred. The goal is to use this insight to better predict what will happen in the future. Data-driven activities should also be utilized to improve existing processes.

- Give analytics an official place in the organizational strategy.
- Expand focus from reporting aggregated data to analyzing underlying behavior.
- Use analytical insight to decide what to do on a strategic level.

Agility

Where agility in the first stage is characterized by its ad hoc nature, this stage will be focused on the formalization of roles and responsibilities. Struggles and conflicts with the old way of working should be smoothed out as much as possible. For every part of the organization someone should be responsible for advocating and integrating the use of analytics and a data-driven mindset. The questions the analytics team is solving are, for a large part, still determined by upper management. They have specific problems that need further analysis. A process should be in place to manage the requests by other departments to the analytics team.

- Formalize roles and responsibilities.
- Assign ambassadorship for analytical initiatives across the organization.

Integration

Further integration takes place in this stage. The analytics team is coming up with new recommendations by digging deeper into the available data. They deliver detailed reports with their findings and recommendations to upper management, whose task it now is to put the changes on the agenda for implementation. Automated reporting and dashboards are getting more sophisticated. Where in the past they had to wait on a report, they can now query a standardized report on demand whenever it suits them.

- Create a process to implement recommendations backed by analytical evidence.
- Create on-demand available reports and dashboards with the right information.

Empowerment

Reports are already providing a way to monitor performance. Managers now might also be interested in improving this performance through a data-driven way of working. However they lack the knowledge and tools to take action. The analytics team should help, educate, and provide tools, to others in the organization that wish to take action. They can help with the implementation of experiments or other activities that can help improve the performance of certain processes. Wide adoption of a data-driven mindset should be nurtured at every opportunity.

- Educate and provide tools to allow managers to use data and create dashboard.

- Nurture the adoption of a data-driven mindset; how can we use data to create success?

7.2.3 Optimizing

Setting up the foundation of analytics is already a great challenge, and will probably lead to great results. But for analytics, this is only the beginning on a long journey to capture the most value out of your data. This contradiction can lead to an organizational struggle, on the one hand you are doing really well, and should continue to go down on this path, on the other hand, you have booked some quick wins and are feeling less sense of urgency to continue to invest in analytical efforts.

The focus in this stage is on developing and operationalizing prescriptive analytics. What should we do in the future to optimize our profitability? Operationalizing insight without manual actions is a technical challenge with great potential for the business and its customers. Recommendation services prescribe products to customers they will more likely respond to.

Data

Large amounts of data are being collected and analyzed. It is now time to bring it to a true big data scale to feed the data-driven models you are creating. Collecting user behavior on websites provides massive amounts of raw data. The key is refining this to user *intent*. What does it mean when someone lingers on this product page? Is he doubting the price or the product model? If you find the intent of a user, you can act on it by personalizing their experience.

- Collect data like user behavior on a bigger scale.
- Use customer data to personalize the experience.
- Focus on data quality.

Metrics

Reports are now standardized and dashboards are available to management. Derived KPI's, backed by analytical models, are proven to be very valuable indicators. The next step is to move this process to real-time operations. Real-time dashboards reveal operational insight, a real-time state of operations. Anomalies are quickly detected and resolved. Continue to show how KPI's evolve over time. It can be a powerful argument for continuing the investments in analytical efforts.

- Define and show operational metrics in real-time.
- Create a process to rapidly detect and deal with defects, anomalies, and trends.

Skills

As analytics continues to evolve, the required skills evolve with it. We move from basic data mining to more advanced concepts, such as machine learning. Machine learning is the science of training a computational model from past experience, to be able to match future events. We can use the large amounts of data we are collecting about customer behavior to attempt to predict future behavior. If we are able to predict future behavior, we can better anticipate the needs of the customer and act accordingly. To be able to do this, we need to have the knowledge to build and train machine learning models. The performance of the models is determined by the quality of the data and the model. A lot of parameters can be tuned, and sometimes intricate knowledge about specific techniques is necessary to build a performing model. The required knowledge and education level depends on how accurate the model has to perform. There are diminishing returns in the amount of effort required to improve the accuracy of a model. Netflix launched a public contest with a $1,000,000 prize to improve their Cinematch algorithm, an algorithm that predicts how a user would rate a movie [56]. Where a trivial algorithm has a error score (root mean square error) of 1.0540, Cinematch already improved this by roughly 10% to 0.9514. It took 3 years and the combined effort of many experts on the subject to improve this score by another 10% to 0.8567. Now for Netflix this improvement can be worth the $1,000,000, but it is important to always closely guard if the return will outweigh the additional effort. Humans do not always act rational, and you do not have perfect information about all the decision factors. It is therefore not always possible to achieve a high prediction accuracy.

- Use new big data sources in machine learning models.
- Use learning analytical models to predict user behavior.
- Use the predicted user behavior to optimize your offerings.

Technology

The technological challenge is to integrate these advanced machine learning algorithms in production systems. There is a massive amount of data and calculation necessary to train these models. Timing is a big challenge: New data sometimes has to lead to insight only moments later. The infrastructure has to be able to quickly transport data. Batch processing has to make way for streaming data. From a pull structure, i.e. the data is queried when needed, to a push structure: When new data is generated, an event, it is pushed and processed immediately. Event-driven architectures are build to react to and handle incoming events. Personalization is a big part of optimizing the customer experience. Automated integration of analytics allows us to adjust the customer journey for every unique individual. This is unfeasible to do by hand for everyone. Machine learning algorithms can prescribe the best journey for every customer. Data collection about the customer is needed to serve as input to the analytical model.

- Invest in technology that can automatically process new emerging data in near real-time.
- Invest in technology to automatically integrate analytical insight into production systems.

Leadership

Leadership has the important task to lead the organization across the chasm from ‘doing analytics’ to ‘breathing analytics’. Great progress has already been made, and the organization might feel it is no longer behind the pack. This may lead to a reduced sense of urgency and less commitment to invest and expand analytical activities. But, we do not want to use analytics to do the same thing, we want to use analytics to make a difference. To make a difference and compete on analytics, you need to continually improve. So this is a critical step for the organization: Are you content with the current situation, or do you go above and beyond?

- Bridge the chasm from doing analytics as a side activity to being a cornerstone of all processes.
- Encourage the organizational culture to be the driving force for analytics, instead of leadership.

Culture

Culture is probably one of the hardest things to change in any organization. Changing culture is a grind, and there is no magic potion. A culture that embraces analytics is necessary to sustain analytical activities. Leadership provides the runway for take-off, but Culture keeps it flying. At this stage a culture of continuous improvement, or *kaizen*, should be stimulated. A culture where everyone is engaged to improve the organization goes hand in hand with our philosophy for an analytical organization. Analytics is born out of the need to break the status quo, and to look for opportunities to improve. Analytics can provide the tools to the organization to analyze the current situation, and improve upon that. Analytics can support a culture of continuous improvement, and that culture can in turn sustain an analytical culture.

- Stimulate a culture of continuous improvement through data.
- Encourage experimentation to test new optimizations.

Strategy

Analytics is continuing to fulfill a more prominent role in the strategy of the organization. We started with analytics in a supporting role in the organization. First as a method to report on current performance, later to further analyze and improve performance. We are now seeing analytics becoming more integral to all processes by augmenting them with analytical insight. In the Reporting and Analyzing stages, we were concerned with how we can we use data to improve the blueprint of a process. In the Optimizing stage and beyond, we are increasingly looking at how can we use analytics to augment the working of a process. The difference is that if you were to stop collecting data in the first scenario, you would still have a better blueprint, say a better performing customer journey, that no longer requires data to further optimize. In the second scenario, the process would break down because you are actively using data, for example to recommend a customer the best product. Analytics has become an active component of the process.

- Position analytics as an integral component of global strategy.
- Use analytics to optimize existing products and processes.
- Use integrated analytics to create an optimized customer experience.

Agility

Analytical processes are getting more interwoven with business processes. Strong execution in the form of collaboration and communication are necessary to create a fluid process. Clear decision and responsibility boundaries should be in place. Continuously rethink if the right people have the responsibility over a certain domain. Agile organizations are able to change fast in part due to a culture that embraces change and innovation. Introducing analytics changes the a business process, so it could be beneficial to change the owner of the process.

- Define or restructure clear decision and responsibility boundaries.
- Embrace change in order to innovate and react faster.

Integration

The Optimizing stage focuses, for a large part, on the automatic integration of analytical insight. More analytics is done on the fly and directly available for use in production systems. Automation means speed, so turning data to insight to actions is getting more rapid. This requires a coordinated effort between analytics and those responsible for production systems. It is not only an analytical challenge as much as a technical challenge.

- Automate the use of analytical insight into existing processes.
- Focus on automatically turning data into actions instead of mere insight.

Empowerment

Employees will be further empowered in two ways. The first is having systems that can track performance over time. While the glaring issues have been tackled in the last stages by analyzing the current situation, the attention now shifts to optimizing the processes. By making small adjustments and tracking results over time, you can steer into the right direction. This is the central philosophy of continuous improvement, and this only works if you can look at the process over a longer period of time. The other improvement comes from the closer integration of analytics into production systems. This allows new data to be reported in real-time. Dashboards can be made that show the current state of operations without any reporting delay. New events can be quickly acted on or resolved before becoming a problem. Because the time, between the occurrence and resolving of an issue, is drastically reduced it becomes easier to also find the source of the issue. Employees should be encouraged to systematically eliminate the sources of these issues, also known as defects in Lean manufacturing.

- Give employees real-time insight into operations and performance.
- Give employees the information to improve their processes.

7.2.4 Empowering

The organization has been successful in continuing their analytical transformation. It is perhaps the easiest route to simply keep doing what you have always been doing. But, there is always room for improvement. This phase focuses on empowering everyone in the organization with the tools, and the knowledge, to analyze and improve their own, or their department's, work. Empowerment can also be in the form of allowing rapid integration of analytical features into products or services. Empowering is an effort that is organization-wide, and hopefully we have already warmed everyone up for change during the previous stages. By continually showing the success of an analytical approach, everyone is more likely to respond positively to change.

Data

Data is already being collected on a massive scale. You probably collect everything there is to collect about your customers. Keep looking out for new opportunities to learn more. Third-party data is another opportunity to get access to valuable data. External companies, specialized in data collection, can sell you market data that complements your own data. Third-party data is often expensive, so make sure it is going to be worthwhile. Data governance is becoming very important if you are starting to grant users, across the organization, access to data. You want employees to have access to all the relevant data, but you also have security concerns. With many access points and a large data volume a clear governance strategy is necessary.

- Look for new data opportunities, including from third-party sources
- Create a clear data governance strategy that scales

Metrics

Instead of focusing on data that is externally oriented, e.g. data about customers or prospects, you can also collect data that is internally oriented, toward your own processes. These metrics can be used to steer the company, and improve the quality of the processes. Quality management concepts, such as Six Sigma, heavily rely on accurate collection of metrics. Amazon uses analytics to perfect their logistics processes, and they collect a lot of data to do so. But, you can also look on how we can collect data to analyze employee performance. Are there predictors for future performance, or if an applicant will be a good hire? Google is known for quantifying their performance management program. The important thing to ask right now is how you can correctly measure these kinds of metrics. Collecting data is one thing, but collecting the right data can be hard. A mix of domain knowledge and statistical knowledge is necessary to come up with the right kind of metrics.

- Collect metrics about internal processes if you have not done so already
- Steer and optimize based on internal metrics
- Look for new opportunities to apply metrics and performance management

Skills

Data science skills will continue to be needed to analyze the data and build analytical models. In addition to this, it is necessary to bring analytical talent to the rest of the organization. Departments, across the organization, should be creating their own roles that analyze and improve the department. They might still rely on a central analytics department to facilitate the technical side, but also start to take on some of the burden themselves. It is beneficial to combine domain knowledge with an analytical background. Business-IT skills come into use, to capitalize on the opportunities of analytics in new products and services. To build a data-driven culture all new hires across the organization can be assessed for their knowledge analytics and data. To get the most out of your data and analytics efforts it is beneficial that the whole workforce is data-literate to some degree.

- Train employees to be more comfortable with basic data analytics

- Democratize and distribute basic analytics across the organization, bring it closer to the work floor
- Bring in the skills to optimally align the business and IT to capitalize on new business opportunities made possible by analytics

Technology

Technological developments in this phase allow business users to do their own form of analytics. You provide the tools and they provide the domain knowledge. Self-service business intelligence is the concept of granting tools to business users to build basic analytics functionality themselves. They can build their own reports or analyze data. The caveat is that these users may not be well versed with data and statistics. It is easy to draw the wrong conclusions when you misinterpret statistics. We are also increasingly building in analytical features into our products and services. Features that add value to the customer by providing them with analytical insight based on the data the product or service collects, e.g. a smart thermometer that learns your preferences.

- Focus on a modular and scalable infrastructure that can rapidly support new opportunities
- Augment products with integrated analytical features (smart products)
- Create more opportunities to use self-service business intelligence

Leadership

The data-driven mindset has now become mature enough to stand on its own legs. The attention of leadership shifts from taking analytics by the hand to encouraging adoption across the organization. Make sure the analytics department delivers on the tools that enable general adoption, and make sure the tools are being utilized. A data-driven mentality is being adopted by the entire organization, this should go all the way to the top. Incentivize and encourage an experimental and analytical approach to work.

- Shift focus from nurturing analytics to encouraging widespread adoption
- Encourage the empowerment of employees and an entrepreneurial spirit

Culture

An analytical culture is slowly becoming a reality. Analytics has found its way to many daily processes. Data and analytics are seen as powerful tools that can be used to accomplish the tasks you are facing. Metrics are used to continuously improve and track progression over time. The sentiment in the organization has shifted from skepticism to optimism. It is important to build on this sentiment to truly become innovative in the future.

- Hire and retain employees with a data-driven mindset
- Build on optimistic sentiment to become a truly data-driven organization

Strategy

The strategy of the organization should strive for the adoption of analytics in all business processes, and new products and services. Data and analytical activities form a central part of the strategy. Data is recognized as a valuable asset that can used or potentially sold. Analytics is used to optimize existing processes, products, and services. New products and services are augmented with 'smart' capabilities that add value for the customer. A strong analytical culture is a major strength that competitors might not have. Continuous improvement and innovation through analytics allow the organization to be competitive. A comprehensive customer view provides a way to better fit products and services to the customer.

- Strive to adopt analytics in every process and product
- Compete with strong analytics and continuous improvement

Agility

With a growth of analytics across the organization, it is important to keep everything in line. Common processes concerning data usage and analytics should become standardized so everyone follows the same rules. Standardized processes are easier to manage and deploy. A standardized plan allows you to capture new opportunities immeditately. Because analytics is not practiced in a central location it is also important to create a platform for sharing expertise. Organizational models, such as that of Spotify, accommodate organization-wide knowledge sharing in so called *guilds*, which consists of workers sharing a specific interest. Keep evaluating if the existing organizational structure is effective or can be improved. Flexibility in organizational structure is valuable to adapt to new opportunities or initiatives.

- Standardize analytics processes and data access in order to deploy in new areas faster
- Share expertise across the organization

Integration

Analytics is spreading everywhere in the organization. From centrally coordinated integration efforts we are moving to analytics being performed closer to the source. Analytics close to the related domain gives the advantage of better domain knowledge and stronger integration. Departments can respond faster to opportunities when they are in charge of their own analytics. Systems with a highly scalable architecture provide easy access to data and analytics, for applications anywhere in the organization. These central hubs allow rapid and easy integration of existing analytical capabilities into new products and services.

- Distribute analysis closer to the relevant departments
- Provide easy integration of analytical capabilities into new offerings

Empowerment

This stage is focused on empowering the organization to engage in analytics. The main way to scale analytics organization-wide is to provide infrastructure centrally and distribute the analysis. Basic analysis can be done with self-service BI tools to create reports and dashboards. Users have access to on-demand reports that they have compiled to their own needs. Advanced analytics can be performed by data scientists with domain knowledge that belong to a specific team or department. Training will empower employees with the knowledge to systematically approach challenges in a data-driven way.

- Scale analytics with self-service BI
- Combine domain knowledge with analytical knowledge to better suit to the needs of departments

7.2.5 Innovating

The final stage is Innovating. Data-driven maturity, and all the things it represents, are on such a level that it is the main source of competitive advantage for the organization. It defines the way business is done, whatever business you are in. This strength opens its doors to many innovative products and services, and can even redefine the business model of the organization. The most important things to realize is that analytical transformation is a journey that never ends. The organization keeps progressing, and keeps redefining itself in everything it does. Continuous improvement is so deeply ingrained in the organization, it becomes second nature.

Data

Data is one of the most valuable assets, if not the most valuable asset in your possession. You also have the capabilities to capitalize on those assets. Because data is so valuable, you are always looking for new way data sources. Companies like Google, that make money through advertising, go to great lengths to build a better profile about their users. Google invests in free software, such as Google Chrome and Android, partially to collect more information that can be used in an advertising profile. Unstructured data is also retrieved from voice and images. Image recognition is already used to make images interpretable to computers.

- Continue to look new data sources from new products
- Leverage alternative unstructured data sources, such as voice and images

Metrics

Metrics are being collected everywhere in the organization, and are being actively used to improve products and processes. KPI's are standardized across the organization. Advanced analytical KPI's are used for strategic goals, and are leading indicators for company performance. They allow organizations to better focus their efforts and measure the success. New KPI's may be introduced that are more in line with the redefined business model of the organization. KPI's have the ability to clearly communicate and incentivize priorities within the organization. Metrics are also being used more broadly to track internal processes and goals, such as employee performance or satisfaction. Metrics found from methods such as cohort analysis can be useful to measure the progress of new ventures. *Innovation accounting* is a technique to measure the success of new ventures instead of relying on traditional accounting methods, which may not be suitable for measuring the success of new innovation.

- Standardize KPI's across the organization

- Track new ventures using innovation accounting methods, instead of traditional metrics

Skills

Mature data-driven organizations house a diverse set of skills. Large organizations, on the leading edge of data science, are often ahead of academic research in fields such as machine and deep learning. They tend to hire academic researchers, with a PhD in computer or data science, for advanced analytics applications. To sustain a culture of data-driven innovation, it is beneficial to have people in the organization with an entrepreneurial mindset and business development skills. Human capital is the greatest asset of an analytical company, because they form the machine that turns data into new business.

- If applicable, employ computer scientists to develop artificial intelligence features
- Develop entrepreneurial en business development skills within the organization to sustain innovation

Technology

Technology is continuously evolving. When you compete on analytics, you are competing on technology. Competitors that are able to better predict customer behavior are more likely to attract and retain customers, while also building a strong relationship that is mutually beneficial. Building and maintaining analytical systems takes a lot of work. We are seeing more software platforms that accelerate the process of building models, allowing data scientists to rapidly churn out new models and reduce the time to market [57]. With automation happening at every step of the data science process, the data scientist can focus on the most important task: asking the right questions.

- Invest in software platforms that accelerate the process of building analytical models
- Focus on asking the right questions

Leadership

Leadership has successfully set up analytics, and has led the cultural reform of the organization. Their task is now to encourage and facilitate innovation based on the powerful analytical capabilities of the organization. New innovation can lead to a business transformation, a fundamental change to how the organization operates. For example, Netflix is known for their strong analytical capabilities, which are used to recommend a user new content. Their prescriptive capabilities have allowed them to build a loyal base of users, because those users were getting personalized recommendations. Now they are taking their analytics one step further, by using the data they collect to create their own content they know will likely succeed. They are slowly transforming, from prescribing the content you like to creating the content you like.

- Stimulate and incubate innovation efforts
- Encourage data-driven innovation; using analytics to determine and capture opportunities

Culture

Culture is the driving force behind data-driven organizations and a sustainable competitive advantage, because it cannot be easily replicated by competitors. Encourage a culture that fosters innovation and a culture of ownership, where employees feel responsibility for the welfare of the organization, without having to be assigned explicit accountability. While a data-driven culture encompasses a lot of factors, the foundation will always be building arguments based on evidence, rather than intuition, and critically looking at every process and activity.

- Encourage innovation and a culture of ownership
- Maintain a data-driven culture, where evidence triumphs intuition

Strategy

Analytics and data become the most important component of the global strategy. Data-driven innovation is the greatest source of competitive advantage. A strong analytical foundation, comprised of data, technology, people, and processes, allows an organization to successfully introduce new services, and compete in new markets. This foundation allows an analytical company like Amazon, traditionally an internet company, to confidently enter a new market with the acquisition of Whole Foods. Amazon has the ability to innovate faster, and make decisions based on evidence backed by data analysis [58]. Expanding into new markets can also be a strategic move, to get access to more data and integrate existing services with new markets. Amazon can use their existing customer insight to provide a better service in a new market, while also having the ability to collect more data; both on a individual customer level, as new customers in general. Because of a network effect combining data sources can lead to better individual customer insight. This effect fuels the pursuit of growth in large internet companies [59]. Data-driven principles are universally applicable and can be internally oriented toward business processes and externally toward consumers and markets. Where we start with the question, 'How can we capture the most value out of our data?', we now ask ourselves: 'How can we generate the most value out of our analytical capabilities?'

- Analytics as a cornerstone of organizational strategy and vision
- Expand into new markets successfully through strong analytical capabilities
- Tailor and optimize the customer experience

Agility

Execution of existing processes is streamlined. Decision domains have been established, the organizational structure is flexible, and suited for knowledge sharing. The organization enables innovation by further developing innovation and change management strategies. These strategies can help breed innovation and successfully bring it to market. The goal is to determine as soon as possible if the new venture is something that sticks, or is a waste of effort. Bring new innovation to the costumers faster and see how they react. Be aware of new opportunities that can be captured quickly due to an agile business structure and culture.

- Manage innovation to increase the success rate of new ventures, and reduce the waste of creating a product nobody needs

- Bring new innovation to market faster to test the waters
- Rapidly capture new opportunities

Integration

We see analytical results further integrated into all business activities. It tells us what to do on a macro-scale: what business ventures to pursue, and on a micro-scale: what do we recommend this customer?.

Analytics moves further to the point where data is generated. For example an IoT devices can already have machine learning capabilities built-in. The advantage is that the data is generated and analyzed on the same device. The computational efforts are also distributed. This architecture allows for massive scaling.

- Use analytics to make all decisions on a strategic scale and on an individual customer scale
- Integrate analytical processing capabilities into all products without centralizing data collection

Empowerment

Everyone has been given access to analytics, and it is now the time to do great things with it. Through organization-wide training initiatives and hiring practices, everyone will have some degree of data-literacy and the right mindset for an analytical organization. Data-driven organizations give employees access to massive amounts of data and analytical capabilities. Combined with the right knowledge and mindset these unlock new opportunities for products and services; either a new product or a new analytical feature. These analytical capabilities also allow you to continuously improve existing processes.

- Give everyone the opportunity to use data and analytical capabilities
- Trust employees with the power to innovate and improve

Chapter 8

Results & Discussion

In this chapter, we will present and comment on the results of the maturity assessment and meta-evaluation. The meta-evaluation reveals what respondents think about the maturity model and assessment. The maturity assessment results are used to analyze the current state of data-driven maturity. The results of the meta-evaluation are used to validate the stages and dimensions of the Data-Driven Maturity Model. The meta-evaluation is also used to validate the questions, accuracy, and viability of the Data-Driven Maturity Assessment. We use the statistical methodologies described in *Section 2.5* of the Methodology chapter.

Section 8.1: Data-Driven Maturity Model, will feature results relevant to our second research question: *What does a data-driven maturity model look like?*. In *Section 8.2: Data-Driven Maturity Assessment*, we dive deeper into the results of the maturity assessment, in order to answer our third research question: *What does a data-driven maturity assessment look like?*.

In addition to the results of the meta-evaluation, we take a deeper look at the outcome of the maturity assessments in *Section 8.3: The State of Data-Driven Maturity.* Are there any observations that can be made from the results of the maturity assessments, and are those observations in line with existing research and observations?

At the end of this chapter, we will discuss any research limitations and additional observations.

8.1 Data-Driven Maturity Model

In this section, we will look at the results relevant to the model composition, in order to determine if the inclusion of all model components is justified and understood. We will analyze all stages and dimensions.

8.1.1 Stages

We will first look at the response to the question *The purpose of this stage is clear.* We analyze this Likert question with a chi-squared test. We expect the purpose of all stages to be statistically clear. The positive and negative responses are grouped under + and -. The neutral responses (○) are not considered in the chi-squared test, because we can not formulate an expectation for the amount of neutral responses. The sample size for the tests with observed neutral responses is effectively the total sample size (17) minus the amount of neutral responses.

H_0: *The observed response is evenly distributed and therefore has a neutral sentiment* H_a: *The observed response is not evenly distributed and therefore has a positive or negative sentiment*								
	Expected		Observed				n = 17	$\alpha = 0.05$
Component	-	+	-	$\circ$	+	Sentiment	p-value	$p \leq \alpha$
Reporting	8.5	8.5	0	0	17	+	< 0.001	true
Analyzing	8.5	8.5	0	0	17	+	< 0.001	true
Optimizing	8.5	8.5	2	0	15	+	0.002	true
Empowering	7.5	7.5	0	2	15	+	< 0.001	true
Innovating	7.5	7.5	1	2	14	+	0.001	true

Table 8.1: Chi-squared test for survey response on stage purpose for all stages.

In Table 8.1, we can see that the purpose of all stages is clear. All stages show a positive sentiment and are significantly different from a neutral sentiment. A clear purpose is an indicator the respondent thinks the addition of the stage makes sense.

The second question we asked is: *How important is this stage in data-driven maturity?* We analyzed the response to this question with a Student's t-test. The answer range is from 0 to 10. For our null hypothesis, we assume a normal distribution around the mid-point of 5. We expect all stages to be significantly important for data-driven maturity.

H_0: *The observed response is equal to or lower than a neutral response* H_a: *The observed response is higher than a neutral response*				
			n = 17	$\alpha = 0.05$
Component	Sample Mean (s)	Sample SD ($\bar{x}$)	p-value	$p \leq \alpha$
Reporting	7.65	2.37	< 0.001	true
Analyzing	8.12	1.22	< 0.001	true
Optimizing	7.88	1.58	< 0.001	true
Empowering	7.88	1.65	< 0.001	true
Innovating	7.65	2.15	< 0.001	true

Table 8.2: t-test for survey response on stage importance for all stages.

In Table 8.2, we can see that every stage importance is significantly different from a neutral distribution. All stages are found to be important for data-driven maturity.

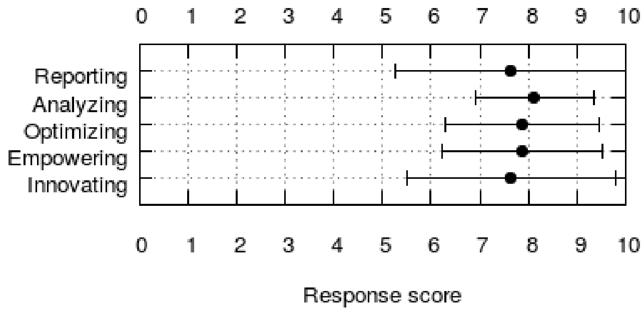

Figure 8.1: Stage importance.

Figure 8.1 shows us the mean and standard deviation of the stages. We can see that the mean of most stages is similar, as is the standard deviation. All stages have a mean above 7.5 and are, therefore, deemed important for data-driven maturity.
The most controversial stage is the Reporting stage with a mean of 7.65 and a standard deviation of 2.37. Participants may perceive that classic business intelligence activities, such as reporting, have no place in data-driven maturity, or should not be distinct stages. The least controversial stage is the Analyzing stage, with a mean of 8.12 and a standard deviation of 1.22. It is interesting to see that two classic business intelligence activities, reporting and analyzing, are respectively the most and least controversial dimensions. Respondents might feel that reporting is not yet 'using data', and analyzing is already using data as an actual input. Respondents might also feel that the Innovating stage is not a primary focus of being data-driven, as data is still used to support processes, instead of enabling them.

8.1.2 Dimensions

We conduct the same analysis for the dimensions of the model. We have asked the same questions as for the model stages. We start with the questions whether the purpose of the dimension is clear. We expect that the purpose for all dimensions is clear.

H_0: The observed response is evenly distributed and therefore has a neutral sentiment *H_a: The observed response is not evenly distributed and therefore has a positive or negative sentiment*									
	Expected		Observed				n = 17	$\alpha = 0.05$
Component	-	+	-	∘	+	Sentiment	p-value	$p \leq \alpha$
Data	7.5	7.5	0	2	15	+	< 0.001	true
Metrics	8.5	8.5	0	0	17	+	< 0.001	true
Skills	7.5	7.5	0	2	15	+	< 0.001	true
Technology	7.5	7.5	0	2	15	+	< 0.001	true
Leadership	8.0	8.0	0	1	16	+	< 0.001	true
Culture	8.5	8.5	0	0	17	+	< 0.001	true
Strategy	8.0	8.0	0	1	16	+	< 0.001	true
Agility	7.0	7.0	0	3	14	+	< 0.001	true
Integration	8.0	8.0	0	1	16	+	< 0.001	true
Empowerment	8.0	8.0	0	1	16	+	< 0.001	true

Table 8.3: Chi-squared test for survey response on dimension purpose for all dimensions.

Table 8.3 shows us the purpose of all dimensions is clear and understood. No dimensions show any significant deviations. A lack of understanding of the purpose of the dimension could indicate the inclusion of the dimension may be unjustified or not understood.

The second question we asked is: *How important is this dimension for data-driven maturity?* We expect our results to be that all dimensions are significantly more important than a neutral distribution.

H_0: *The observed response is equal to or lower than a neutral response* H_a: *The observed response is higher than a neutral response*				
			n = 17	$\alpha = 0.05$
Component	Sample Mean (s)	Sample SD ($\bar{x}$)	p-value	$p \leq \alpha$
Data	7.59	2.24	0.003	true
Metrics	7.76	1.20	< 0.001	true
Skills	8.53	1.23	< 0.001	true
Technology	7.76	1.52	< 0.001	true
Leadership	8.18	0.95	< 0.001	true
Culture	8.00	1.17	< 0.001	true
Strategy	8.18	1.13	< 0.001	true
Agility	7.12	1.41	< 0.001	true
Integration	8.00	1.00	< 0.001	true
Empowerment	7.65	1.17	< 0.001	true

Table 8.4: t-test for survey response on dimension importance for all dimensions.

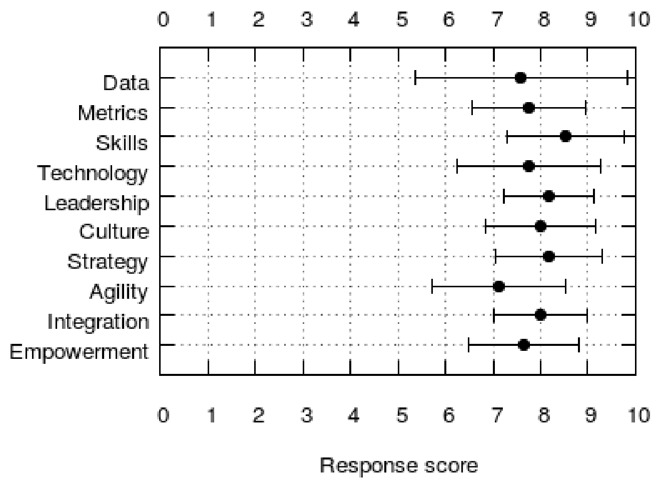

Figure 8.2: Dimension importance.

The statistical analysis in Table 8.4 shows us some variance in the importance of model dimensions. Figure 8.2 shows the mean and standard deviation of all dimensions. One of the most controversial dimension is quite surprisingly the Data dimension, with a standard deviation of 2.29. We will revisit and discuss this result in depth in Section 8.5.1. The Agility dimension has the lowest mean, with a score of 7.12.
Skills is found to be the most important dimension. Organizations seem to recognize the vital importance of having not only the technology but also the skills to excel in analytics. Two of the

least controversial dimension are Leadership and Strategy. Respondents agree on the importance of leadership and steering activities from the top.

We cannot conclude that other potential dimensions or stages would have also fit the model. We can only validate the chosen components in our model and conclude that our representation of data-driven maturity is a validated view. We have not received any feedback suggesting other aspects or dimensions that or not covered by this model.

8.2 Data-Driven Maturity Assessment

The Data-Driven Maturity Assessment forms a crucial part in bridging the gap between the artifact, the Data-Driven Maturity Model, and practice. The attribute-questions in the assessment should be relatively easy to answer for a broad audience. Intricate technical knowledge should not be a necessity. Yet the attributes should accurately reflect the most important aspects of the dimension. An accurate maturity indication is necessary to provide the most relevant information to the assessor. We tested the validity, accessibility, accuracy, viability, and usability of the maturity assessment.

8.2.1 Validity

Validity means all the questions asked are relevant to data-driven maturity. We will look at the response to the question *How important is this attribute for data-driven maturity?*. We expect all attributes to be important for data-driven maturity. The response is analyzed using a t-test. The results, covering the next two pages, are in Tables 8.5 and 8.6.

H_0: *The observed response is equal to or lower than a neutral response* H_a: *The observed response is higher than a neutral response*				
			n = 17	$\alpha = 0.05$
	Sample Mean (s)	Sample SD ($\bar{x}$)	p-value	$p \leq \alpha$
Data				
Analytics Data Storage	7.12	2.47	< 0.001	true
Source Variety	6.53	2.37	< 0.001	true
Governance	7.65	1.97	< 0.001	true
Metrics				
Sophistication	8.41	1.23	< 0.001	true
Timeliness	7.59	1.46	< 0.001	true
Usage	7.82	1.24	< 0.001	true
Skills				
Analytical Skills	8.35	1.06	< 0.001	true
Focus	8.82	0.95	< 0.001	true
General Skills	7.65	2.06	0.001	true
Technology				
Infrastructure	8.06	1.75	< 0.001	true
Capabilities	7.82	1.29	< 0.001	true
Leadership				
Activities	8.65	1.17	< 0.001	true
Attitude	8.24	0.97	< 0.001	true

Table 8.5: t-test for survey response on attribute importance for Data, Metrics, Skills, Technology, and Leadership.

	Sample Mean (s)	Sample SD ($\bar{x}$)	n = 17 p-value	$\alpha = 0.05$ $p \leq \alpha$
Culture				
Attitude	7.82	1.67	< 0.001	true
Adoption	7.82	1.59	< 0.001	true
Strategy				
Role	8.12	0.93	< 0.001	true
Focus	7.94	0.97	< 0.001	true
Agility				
Process Maturity	7.41	1.06	< 0.001	true
Roles & Responsibilities	6.76	1.95	0.002	true
Integration				
Method of Integration	8.24	1.15	< 0.001	true
Action Capabilities	7.94	0.97	< 0.001	true
Empowerment				
Tools	7.71	1.36	< 0.001	true
Education	7.76	1.15	< 0.001	true
Information	7.71	1.21	< 0.001	true

Table 8.6: t-test for survey response on attribute importance for Culture, Strategy, Agility, Integration, and Empowerment.

All attributes pass validation, no superfluous attributes are included in the maturity assessment. The attributes of the already contested Data dimension are the most controversial with standard deviations of 2.47, 2.37, and 1.97.

8.2.2 Accessibility

To test the accessibility of our maturity assessment, we added the following the Likert question *I have enough knowledge about the organization to answer this question.* The 'question' refers to the maturity question about the attribute. We expect that most respondents feel knowledgeable enough to answer the questions.

H_0: The observed response is evenly distributed and therefore has a neutral sentiment

H_a: The observed response is not evenly distributed and therefore has a positive or negative sentiment

	Expected		Observed				n = 17	$\alpha = 0.05$
Attribute	-	+	-	○	+	Sentiment	p-value	$p \leq \alpha$
Data								
Analytics Data Storage	8.0	8.0	0	1	16	+	< 0.001	true
Source Variety	7.5	7.5	0	2	15	+	< 0.001	true
Governance	4.0	4.0	0	2	15	+	< 0.001	true
Metrics								
Sophistication	6.5	6.5	0	4	13	+	< 0.001	true
Timeliness	7.5	7.5	0	2	15	+	< 0.001	true
Usage	8.0	8.0	0	1	16	+	< 0.001	true
Skills								
Analytical Skills	7.5	7.5	0	2	15	+	< 0.001	true
Focus	8.0	8.0	0	1	16	+	< 0.001	true
General Skills	5.5	5.5	0	6	11	+	0.001	true
Technology								
Infrastructure	8.0	8.0	0	1	16	+	< 0.001	true
Capabilities	8.0	8.0	0	1	16	+	< 0.001	true
Leadership								
Activities	8.5	8.5	0	0	17	+	< 0.001	true
Attitude	8.5	8.5	0	0	17	+	< 0.001	true

Table 8.7: Chi-squared test for survey response on attribute knowledge for Data, Metrics, Skills, Technology, and Leadership.

	Expected		Observed				n = 17	$\alpha = 0.05$
Attribute	-	+	-	○	+	Sentiment	p-value	$p \leq \alpha$
Culture								
Attitude	8.0	8.0	1	1	15	+	< 0.001	true
Adoption	8.5	8.5	1	0	16	+	< 0.001	true
Strategy								
Role	8.0	8.0	0	1	16	+	< 0.001	true
Focus	7.5	7.5	0	2	15	+	< 0.001	true
Agility								
Process Maturity	6.5	6.5	0	4	13	+	< 0.001	true
Roles & Responsibilities	6.5	6.5	1	4	12	+	0.002	true
Integration								
Method of Integration	8.0	8.0	0	1	16	+	< 0.001	true
Action Capabilities	8.0	8.0	0	1	16	+	< 0.001	true
Empowerment								
Tools	8.0	8.0	0	1	16	+	< 0.001	true
Education	7.5	7.5	0	2	15	+	< 0.001	true
Information	8.0	8.0	0	1	16	+	< 0.001	true

Table 8.8: Chi-squared test for survey response on attribute knowledge for Culture, Strategy, Agility, Integration, and Empowerment.

All attributes have a positive sentiment on the knowledge question. The roles of the respondents were diverse, roles included: business analyst, innovation manager, head of IT strategy, project manager, CEO, CTO, information manager, and BI manager. Most roles do not require a hard technical background, but almost all were still able to answer the questions with confidence. This satisfies one of the objectives of the maturity assessment: to provide an accessible way of assessing data-driven maturity without intricate knowledge of the model or extensive technical knowledge.

8.2.3 Accuracy

We will analyze if the output of the scoring model is accurate. After the assessment respondents were served a report. An example report can be viewed in Appendix 11.1. This report contains the scores for each dimension and an accompanying text with information from the maturity model, based on the current stage. We asked *How accurate is the maturity score for this dimension?*, on a 0 to 10 scale. We also asked for the accuracy of the overall maturity score, which is an average of the maturity scores for every dimension.

H_0: The observed response is equal to or lower than a neutral response

H_a: The observed response is higher than a neutral response

			n = 17	$\alpha = 0.05$
	Sample Mean (s)	Sample SD ($\bar{x}$)	p-value	$p \leq \alpha$
Data	7.53	1.12	< 0.001	true
Metrics	7.88	1.36	< 0.001	true
Skills	7.29	1.57	< 0.001	true
Technology	8.00	1.32	< 0.001	true
Leadership	7.71	1.21	< 0.001	true
Culture	8.12	1.05	< 0.001	true
Strategy	8.06	0.97	< 0.001	true
Agility	7.24	1.71	< 0.001	true
Integration	7.65	1.22	< 0.001	true
Empowerment	7.24	1.03	< 0.001	true
Overall	7.24	1.35	< 0.001	true

Table 8.9: t-test for survey response on report accuracy for all dimensions and overall

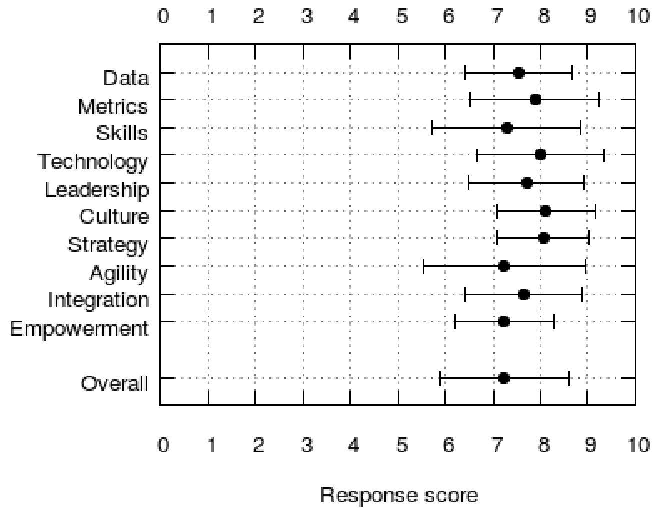

Figure 8.3: The accuracy of report scores for all dimensions and overall score.

All dimensions show a statistically significant accuracy, and score means above 7 out of 10. The asked questions are sufficient to accurately estimate the data-driven maturity of a dimension. Interesting to note: We can see in Figure 8.3 that the overall accuracy is rated lower than the accuracy for individual dimensions. Respondents might feel that not all facets of data-driven maturity are represented. Our scoring model evenly weighs all dimensions for the overall score. For simplicity, but also to not promote leaving one dimension behind the rest in importance. Respondents might feel that certain dimensions should weigh more than other dimensions in the calculation of overall maturity.

8.2.4 Viability

To test the viability of the maturity assessment, we looked at if the outcome of the maturity assessment, the personalized report, can help the organization in becoming more data-driven. Before the report, we asked the following three Likert questions: *My organization is planning to become data-driven* (intent), *We have a clear idea on how to become data-driven* (strategy), *We can use more information on how to become data-driven* (information). After the report, we asked the following three Likert questions: *The purpose of the Data-Driven Maturity Assessment is clear* (purpose), *The scope of the Data-Driven Maturity Model and Assessment are clear* (scope), and *The results of the Data-Driven Maturity Assessment can help us create a strategic plan for data-driven maturity* (results).

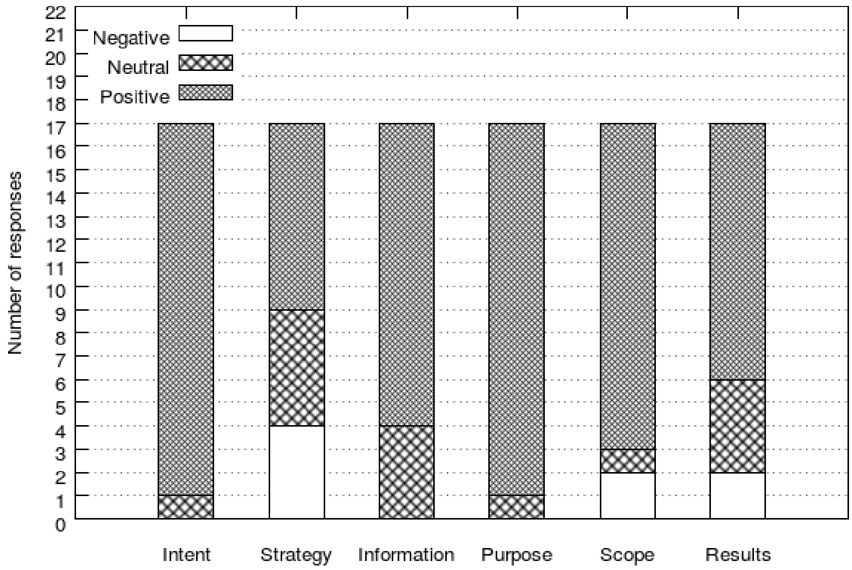

Figure 8.4: The survey response on intent, strategy, information, purpose, scope, and results

H_0: The observed response is evenly distributed and therefore has a neutral sentiment

H_a: The observed response is not evenly distributed and therefore has a positive or negative sentiment

	Expected		Observed				n = 17	$\alpha = 0.05$
Question	-	+	-	○	+	Sentiment	p-value	$p \leq \alpha$
Purpose	8.0	8.0	0	1	16	+	< 0.001	true
Scope	8.0	8.0	2	1	14	+	0.003	true
Results	6.5	6.5	2	4	11	+	0.013	true

Table 8.10: Chi-squared test for survey response on purpose, scope, and results

The response on purpose, scope, and results is in Table 8.10 and Figure 8.4. We can immediately see that almost all organizations have the intent to become (more) data-driven. However, half of the assessed organizations do not have a clear idea on how to become data-driven. The intent is there, but the strategy is missing. Many organizations are still struggling to get started. With this in mind, it comes as no surprise that almost all organization can use more information on how to become data-driven. There is a clear gap in knowledge, and how to bring this knowledge to practice.

The purpose and scope of the assessment are clear: The response is positive and significantly different from a neutral response. The response to whether the results can help in creating a strategic plan is more varied. 2 respondents did not think the results could help, 4 respondents were unsure, and 11 respondents agreed that the report could help in creating a strategic plan. The majority of the response is positive that the maturity report can help in creating a strategic plan.

It is clear that organizations do have the intent to become data-driven and can use help in doing so. The scope and purpose of the maturity model are clear. The results have shown to be useful in the majority of cases. We have to acknowledge that respondents were actively approached to complete the assessment, and may not have been looking for this type of feedback and information. For some organizations, it is a useful tool to gauge the current maturity level and to measure the progress along the way. The information in the full maturity model can also prove to be a valuable source of information, but we are unable to quantitatively test the usefulness of the full 20+ page maturity model. Additional feedback mentioned that the assessment helped to generate an overview of data-driven maturity, and provided a reflection on the current state of the organization. Some feedback mentioned that a single assessment for the entire organization proved to be difficult and ineffective.

8.2.5 Usability

After resolving many issues during the first iterations of our maturity assessment, the last respondents completed the assessment without any usability issues. The chat for assistance was not consulted during the last phase of development. The assessment is accessible online and works on the majority of web browsers.

Most respondents completed the assessment in 30 to 60 minutes. We were unable to test how long the assessment would take if the meta-questions were omitted. It is fair to assume the assessment can be done in 30 minutes.

8.3 The State of Data-Driven Maturity

In addition to the results of the meta-evaluation, we also look at the outcome of the maturity assessments by the participating organizations. Our sample size is insufficient to draw conclusions about the entire population of organizations in, for example, the Netherlands. However, we have managed to assess 14 different organizations, in many different industries. We will look at any interesting observations emerging from our results.

We will only analyze the collective result, and not take an in-depth look into individual results. Our interest is into any observations in the maturity levels of the assessed organizations.

We have plotted the average maturity score and standard deviation of every dimension. The standard deviation is the standard deviation for all the maturity scores for one dimension, between all maturity assessments. A high standard deviation might mean that organizations are on very different levels for this dimension. It might be that some organizations have a strong analytical culture that is ahead of the other dimensions, while other organizations are lagging behind in the Culture dimension.

Are organizations already very data-driven, or are they still in the beginning stages of data-driven maturity? McKinsey estimates that many organizations are still not capturing the most value out of their data [1]. This research suggests that many organizations are still in the beginning stages of data-driven maturity.

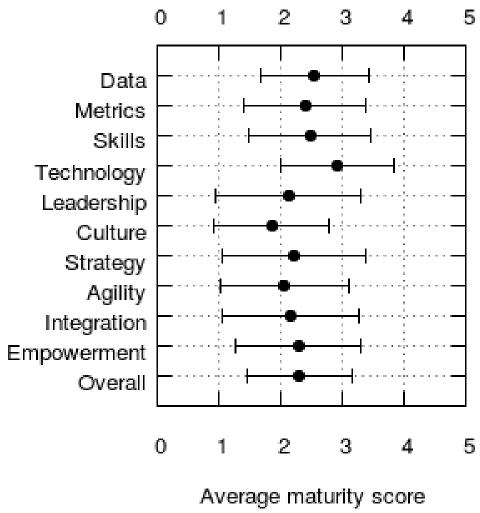

Figure 8.5: Average maturity score and standard deviation for every dimension.

As we can see in Figure 8.5, the average dimension scores do not drift too far apart. This indicates that there are no dimensions on which every organization scores much higher than another dimension. The lowest average dimension score is for the Culture dimension, with a score of 1.85. The highest average dimension score is 2.91 for Technology.

The standard deviations for all dimensions are all very consistent, ranging from 0.88 for Data to 1.17 for Strategy. The Strategy dimensions shows the biggest discrepancy between organizations. Some organizations may have a more heavy focus on data and analytics, while others are lacking an official strategy.

As seen in Figure 8.5, the overall average assessment score is 2.30, indicating the first steps in the Optimizing stage. We cannot generalize our findings, due to our limited sample, but it seems that the majority of organizations are still in the beginning stages of data-driven maturity. They have the intent, are able to do basic analysis and reporting, but are struggling to become more data-driven.

In Figure 8.6 below, the distribution of overall maturity scores is showed in a histogram, with a bin size of 0.5, or half the width of a stage.

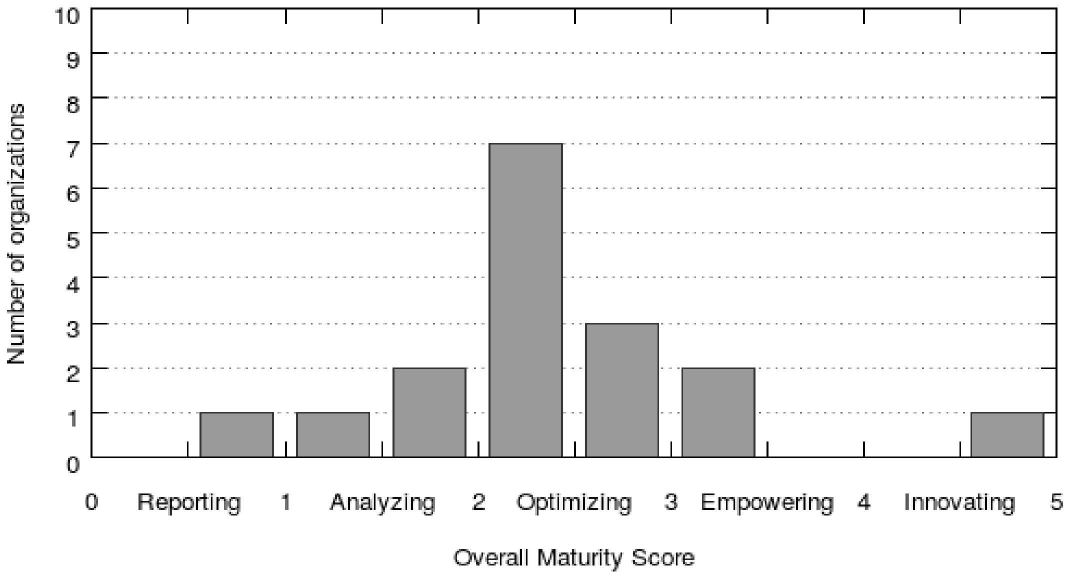

Figure 8.6: Distribution of overall maturity scores.

In the overview in Figure 8.6, we can see the maturity scores for every assessed organization. It starts at 0, indicating no maturity, and progresses to 5, indicating full maturity at an Innovating level. The majority of the organizations are in the Analyzing and Optimizing stages, between 1 and 3. Most organizations have just achieved an Analyzing level of maturity, and are making their first steps into the Optimizing stage.

We cannot draw any conclusions about data-driven maturity for an entire population, outside of our sample. Our results seem to indicate that a lot of organizations have not matured far beyond basic reporting and analytics capabilities. We believe data-driven maturity can truly shine when it matures beyond reporting and analysis, by further integrating analytical insight into business processes and products.

8.3.1 From Data Science to Data-Driven

Two interesting observations can be made in the maturity score results. The first is that the majority of organizations have basic reporting and analytics capabilities. However, not many organizations have progressed beyond these basic capabilities. The capabilities in the first two stages are business intelligence capabilities. The study by McKinsey suggests that many organizations still have a lot of value to capture from their data. Our results support that notion, the majority of organizations have not yet entered the final stages, the stages where an organization can truly differentiate itself from its competitors.

The other observation can be made in Figure 8.5. There is a discrepancy between the first four dimensions and the other six, the first four having higher scores on average. We have to note that this discrepancy can be the result of a design flaw in the maturity assessment. However, looking back at the accuracy results in Section 8.2.3, respondents did agree the dimension scores were accurate. What the first four dimensions - Data, Metrics, Skills, and Technology - have in common, is that they are all focused on data science, and they are some of the most tangible components of data-driven maturity. In another McKinsey article [60], one of the insights is that "embedding analytics is as much about change management as it is about data science". It suggests that the solution is not to simply 'hire technology people and upgrade software', but a strong change management strategy is needed. Our results indicate that many organizations are ahead in the data science components, strong Data and Technology, but are lacking in the organizational components, such as Culture and Integration. It is telling that Culture and Technology, a soft organizational dimension, versus a hard technical dimension, are respectively the lowest and highest scoring dimensions.

8.4 Research Limitations

In this section, we will discuss and comment on the limitations of the conducted research.

8.4.1 Sample Selection

Gathering study subjects for our assessment was a challenge. For a quantitative analysis of the results, we have to study a sufficient sample size. However, the assessment posed two major challenges for potential respondents. The effort required for completing the assessment was high: The assessment took between 30 and 45 minutes to complete on average. A 5-minute survey would have attracted more respondents. Potential respondents are more likely to invest 5 minutes to help the researcher. A 45-minute assessment is less likely to passively attract respondents. We were aware of this from the start. We aimed to overcome this by making the assessment as short and smooth as possible. We aimed to create a careful balance between comprehensiveness and speed. This also meant we had to limit the questions we could ask and the amount of data we collected. These efforts led to a more compact assessment which can be completed even faster if the meta-questions do not have to be answered. Unfortunately, we could not further reduce the time it takes to complete the assessment without further compromises. We made the assessment as practical and rewarding as possible by instantly providing a personalized report after the assessment to incentivize potential respondents.

The subject of the assessment is also a potential limitation for respondents. Data-driven maturity and analytics are a subject not all people have extensive knowledge of. This greatly reduces the number of potential candidates. We aimed to make the assessment as accessible as possible, without the need for extensive technical knowledge. We succeeded in creating an assessment for which the majority of the respondents had sufficient knowledge to answer all the maturity questions. However potential candidates may have still felt this assessment was not useful or applicable. Because of the required work and suitability of the potential participants, passive acquisition through LinkedIn posts was ineffective. This meant we had to rely on active acquisition through our network relations. This opened us up for some bias in our sample selection. For example, the approached organizations are more inclined to participate out of a favor, instead of active interest or need. We also needed to ensure we sampled organizations outside of our own industry. We have succeeded in this, assessed industries include: finance, e-commerce, transport, and healthcare. Assessed organizations also range in size from small (< 50 employees) to large (> 1000 employees). All assessed organizations were based in the Netherlands. We were aware we would be unable to sufficiently sample outside of the Netherlands.

8.4.2 No Longitudinal Observations

Due to the nature of our study, we were unable to make longitudinal observations. Reaching full data-driven maturity, as defined in our model, requires a large effort over multiple years. To be able to test the full effectiveness, we would have to conduct a longitudinal study over multiple years, which would be impossible in our time span. This limits our ability to study the practical effectiveness of the model and assessment. We can only make conclusions whether the outcome has the potential to start activities or help with initial activities. Future case study research could include longitudinal studies of the effectiveness of the maturity model.

8.5 Further Discussion & Additional Observations

8.5.1 Data Dimension

The Data dimension is one the most controversial dimensions. The sample mean is 7.59 and the sample standard deviation is 2.24. Respondents are divided on the importance of the volume and variety of data. We have already excluded the volume attribute from the final assessment. Data is controversial because it is very situational to the organization in what data can be collected and what data is relevant. For example, the importance of data volume cannot be generalized to all organizations, some organizations may benefit more from the rapid processing of smaller amounts of data. However, organizations can still benefit from expanding their horizons beyond what they think might be relevant. Part of the allure of big data is finding insight where you might not expect to find insight. Autonomous analytics can work on huge data sets to reveal novel insight that would have been impossible to find by hand. We do agree with the notion that applying the insight from data is more important than simply collecting large volumes of data.

8.5.2 Local versus Global Scope

One of the most interesting things we have heard is that it is hard to gauge the maturity of the entire organization. Respondents found it hard to answer some questions because some parts of the organization are on a completely different level than others. We have introduced a 'scope' field to the assessment, so the respondent can scope their response to that particular department or business unit. The different maturity levels within the organization are interesting because we expected units within an organization to have similar levels of maturity. Differences suggest that best practices are not shared and there is a lack of central coordination or global strategy. Cross-sectional assessments of multiple divisions in the organization could help to identify discrepancies in the organization.

8.5.3 Theory Formalization

The results show that, in general, the respondents agreed on the components of data-driven maturity. No additional components were suggested. One of the objectives of this research was to formalize some of the known best practices and theory of data-driven maturity and analytics. In this regard, it seems our model covers most aspects. We have also heard the model and assessment 'confirm what I already know'. The model successfully formalizes theory from grey literature.

Chapter 9

Recommendations & Applications

The main objective of this research is to help bridge the gap between theory and bringing analytical capabilities to practice. In this chapter, we discuss key takeaways and potential applications of this research and the accompanying Data-Driven Maturity Model and Assessment. This chapter is meant for practitioners in the field. We dive deeper into how the model and assessment can be used as a valuable tool for organizations.

9.1 Introduction

Analytics and data-driven maturity are subjects that are hard to grasp. The possibilities seem endless, but no action is taken. A maturity model, a framework of knowledge and best practices, progressing from incapability to full capability, can help to organize the vast amount of information. Combined with an accessible maturity assessment it becomes a powerful tool that can help to assess the current maturity level of the organization and select the most relevant parts out of the maturity model. We will discuss some potential applications in the following sections.

9.2 As a Standalone Maturity Assessment

The most straightforward way to apply the maturity assessment is to have one person in the organization, preferably someone on a management level and with sufficient knowledge about analytics (CIO, CTO), take the assessment. The participation in the assessment already provides a wealth of information. The outcome of the assessment can be used to determine the current competencies and weaknesses. With the help of the action points in the report, the foundation of a strategic plan can be formed. The assessment determines the as-is situation of the organization. The report can help in creating a plan for achieving a to-be situation. The assessment and report cannot substitute for the intricate knowledge about the challenges and opportunities in the organization that an assessor might have. Combine the best practices from the maturity model with the local opportunities to achieve the best of both worlds.

9.3 As a Data-Driven Maturity Framework & Blueprint

The data-driven maturity model provides a strong conceptual overview of all facets of data-driven maturity. Strategists and educators can use this maturity model as a conceptual framework of data-driven maturity. It can be used to communicate or teach concepts related to data-driven maturity. Dividing or categorizing plans into 10 dimensions can help to create order and clarity. The concepts addressed in the model can also be used as arguments for change. Plans and arguments are stronger when backed by a solid conceptual background.

9.4 As a Cross-Sectional Assessment

One of the emerging applications of the data-driven maturity assessment is to conduct a cross-sectional assessment of the organization. Multiple assessments can be combined to create a more accurate result or as a starting point for discussion. A cross-sectional assessment can be conducted in multiple ways. The first and most simple strategy is to have the assessment taken by multiple persons, all with sufficient knowledge, and to discuss the results. Discrepancies between assessments can raise points for discussion. A 360-degree view on the subject can provide valuable insight. The other strategy, applicable for larger organizations with multiple divisions or large departments, is the compare the results of multiple divisions within the organization. We have seen in larger organizations that the results for one organization can differ greatly by division. Valuable knowledge is not transferring to other divisions. Communication of best practices within the organization can help to bring lagging divisions up to par.

9.5 As a Measure of Progress

The data-driven maturity assessment can be used to measure the progress of the organization over a longer time. For example, an assessment can be taken every quarter. The report of that assessment can be compared to the previous quarter. What projects are progressing well and which departments are struggling to make progress? However, we should be aware that, as with all self-reported studies, the results of the maturity assessment can be easily inflated. Take caution to not incentivize a higher report score. You can also apply a cross-sectional assessment to average the results.

9.6 Summary

The data-driven maturity model and assessment can prove to be a useful tool for any organization. In addition to the information found in the model, the assessment can be utilized in multiple ways to assess maturity, discuss data-driven concepts, and measure the progress the organization is making.

Chapter 10

Conclusions & Future Work

In this chapter, we summarize our results and answer our research questions. We also explain the scientific value of our findings and the potential direction for future work in the final sections.

10.1 What is a data-driven organization?

Our first research question, *What is a data-driven organization?*, is divided into two sub-questions. *What is the difference between business intelligence & analytics and data-driven?* and *What are the characteristics of a data-driven organization?*.

We have taken an extensive look at different maturity models in related domains in Chapter 3. *Data-driven* is a relatively new term with a lack of scientific work. The term *data-driven* simply means driving decisions and actions with data. Business intelligence & analytics (*analytics*) is a data-driven activity and a major part of being data-driven. But we have also seen that data-driven goes further than analytics. Where analytics is often an isolated activity within the organization, data-driven extends toward the entire organization. We have reflected this in our maturity model with the inclusion of two novel dimensions: Integration & Empowerment, and two dedicated stages: Empowerment & Innovation. These components are perceived by our respondents ($n = 17$) as valuable additions to the maturity model.

The analysis of the data-driven maturity levels in Section 8.3 supports the insight that data-driven organizations are more than just data science. Our sample is too limited to draw definitive conclusions, but the assessed organizations score higher on the classic analytics dimensions than on the organizational dimensions. The early focus in many organizations seems to be on the data science aspects of data-driven maturity.

Through our literature review of existing models and feedback of experts, the characteristics of data-driven organizations can be identified. This is in addition to the dimensions of our data-driven maturity model, which we will discuss in the next section.

Data-driven organizations can be best characterized by their desire to turn data into action and their organizational approach. The desire and ability to turn data into action is a challenge with many components. Data needs to be sourced, processed, analyzed, and turned into a decision as fast as possible. This is best reflected in our Data, Skills, and Technology dimensions. The other side is the organizational approach of data-driven organizations. Organizations that successfully capture the value of data do not do so by isolating their data activities. It is an organization-wide effort, which is why we have put extra emphasis on organizational components instead of purely technical or analytical components. Dimensions such as Leadership and Culture are not always represented in existing maturity models but are an important factor in reaching data-driven maturity. This organizational approach is best reflected in our Leadership, Culture, Strategy, Integration, and Empowerment dimensions.

10.2 What does a data-driven maturity model look like?

The second research questions is *What does a data-driven maturity model look like?* with 2 major components: *What stages does a data-driven maturity model have?* and *What dimensions does a data-driven maturity model have?*

An analysis of existing maturity models in Section 3.3 showed the need for a novel maturity model. We can conclude that the Data-Driven Maturity Model fills the gaps in the current landscape of maturity models. The model fulfills the points raised at the end of Section 3.3: formally built, validated by practitioners, and has a documented methodology and design process. The model also covers emerging thoughts on data-driven maturity: starting with identifying information needs, emphasizes the importance of culture and sponsorship, and integrating analytical insight into business processes.

The results in Section 8.2.4 show us that 16 out of 17 respondents said their organization has the intention to become more data-driven. 13 out of 17 respondents agreed that they can use more information on how to become data-driven. This indicates there is a need for a Data-Driven Maturity Model. The stages and dimensions pass our statistical validation. They are tested on their importance for data-driven maturity. None of the components were rejected by the respondents.

Stages

- Reporting: Visualize existing data and create the foundation for an analytical future.
- Analyzing: Dive deeper into the data to achieve insight into why things happened.
- Optimizing: Optimize business processes by integrating analytical insight.
- Empowering: Empower employees and products with data and analytical capabilities.
- Innovating: Use data and experiments to innovate in products and transform the organization.

Dimensions

- Data: The fuel for all data-driven activities. How do you source and manage your data?
- Metrics: The key to measuring output and managing performance. How do you use, collect, and enrich your KPI's?
- Skills: Essential for operating a data-driven organization. Do you hire and educate the right people?
- Technology: The foundation for a data-driven organization. What technology do you need to build an analytical process?

- Leadership: The cornerstone for a successful analytical transformation. How does leadership successfully steer the transformation?
- Culture: The driving force behind a data-driven organization. How does culture affect and promote data-driven adoption?
- Strategy: The plan for success. What role does analytics have in your plans and vision of the future?
- Agility: The ability to adapt and succeed. How well are your roles and processes organized to change and deliver?
- Integration: The integration of analytical insight into processes. How is the organization using analytical output and how are operational processes using real-time intelligence?
- Empowerment: The empowerment of the organization. How is data analytics helping your employees to succeed?

10.3 What does a data-driven maturity assessment look like?

In this section, we look at our third research question: *What does a data-driven maturity assessment look like?* The question has two sub-components. *What questions does a data-driven maturity assessment have?* and *Can the outcome of the data-driven maturity assessment help organizations in settings steps toward data-driven maturity?*.

We conclude that the Data-Driven Maturity Assessment has proven to be a valuable tool. We have proven the validity and accessibility of all maturity questions. The results in sections 8.2.2 and 8.2.1 show us that the questions in the assessment are found to be both important, and do not require extensive knowledge to answer.

The accuracy of the assessment is proven, the overall maturity score has a 7.24 out of 10 on accuracy. All dimensions are found to be accurate, with accuracy scores from 7.24 to 8.06.

The viability of the model is found to be high. The results in Section 8.2.4 show that the majority of the organizations agree that the results of the maturity assessment can help to create a strategic plan. The purpose and scope of the assessment are found to be clear. About half of the assessed organizations did not have a clear idea on how to become data-driven. The Data-Driven Maturity Assessment can help those organizations to create a plan for action. The action points in the report provide advice to implement. We conclude that the Data-Driven Maturity Assessment is a viable strategic tool that can be used in a broad spectrum of organizations.

Questions on the following subjects are validated on their importance, provide an accurate estimate of data-driven maturity, and are therefore included in the final assessment.

Data

- Analytics Data Storage
- Source Variety
- Governance

Metrics

- Sophistication
- Timeliness
- Usage

Skills

- Analytical Skills
- Focus
- General Skills

Leadership

- Activities
- Attitude

Strategy

- Role
- Focus

Integration

- Method of Integration
- Action Capabilities

Technology

- Infrastructure
- Capabilities

Culture

- Attitude
- Adoption

Agility

- Process Maturity
- Roles & Responsibilities

Empowerment

- Tools
- Education
- Information

10.4 Research Relevance

Applying better decisions to business challenges is of tremendous economic value. Some organizations are already reaping the full benefits by applying data-driven decision mechanisms, on all levels, and across the organization. Many organizations are eager to adapt to this new world because of the huge potential and the fear of lagging behind. These organizations do have the intent but lack the skills and knowledge to reach data-driven maturity. Our Data-Driven Maturity Model and Assessment help to close this gap in skill and knowledge by providing easy to use tools that help organizations in determining their current maturity level and provide an overview of all relevant components. The maturity report provides the most relevant information to the current challenges.

The research domain is currently still underdeveloped. Analytics and data-driven maturity are concepts that, although rooted in older theory, rose to prominence fairly recently. The organizational aspect of being a data-driven organization lacks a formal scientific body of work. Best practices are shared in grey literature and are based on anecdotal evidence. The resulting maturity models in this domain are not developed with a formalized methodology. One objective was to digest this grey literature, combine the results, and validate the result. With an iterative development methodology, we were able to formalize the best practices and known theory. The resulting maturity model can now function as a scientific artifact for future research on data-driven maturity. This research also took the steps to put analytics as an activity in an organizational perspective, as part of the whole.

10.5 Future Work

The data-driven maturity model and assessment focus on starting a journey toward data-driven maturity and providing the necessary strategic information. As discussed in *Section 8.4: Research Limitations*, an in-depth look into the actual journey could not be realized. Future research could focus on the main challenges or success factors during this journey toward full maturity. This research can be conducted through a case study or observing organizations that are already more mature. Future research questions could include *What are the key success factors for reaching data-driven maturity?* or *What are the main pitfalls in reaching data-driven maturity?*. The results of this future research can be further incorporated in a future version of the Data-Driven Maturity Model.

The maturity assessment has also proved to be an accurate assessor of current data-driven maturity. The assessment can therefore be used as an assessment protocol for a larger sample of organizations. In the future, we can assess and compare the state of data-driven maturity on a larger scale. We could compare organizations in different countries on their level of data-driven maturity. One interesting direction is the discrepancy found between data science dimensions and organizational dimensions. The assessed organizations scored lower on the organizational dimensions. Future research can look into this for a larger population.

Bibliography

[1] Nicolaus Henke, Jacques Bughin, Michael Chui, James Manyika, Tamim Saleh, Bill Wiseman, and Guru Sethupathy. The age of analytics: Competing in a data-driven world. *McKinsey Global Institute*, 2016.

[2] Daniel J. Power. A brief history of decision support systems. `http://dssresources.com/`, 2017.

[3] George A. Gorry and Michael S.S. Morton. A framework for management information systems. *Sloan Management Review*, 1971.

[4] Hans P. Luhn. A business intelligence system. *IBM Journal of Research and Development*, 2, 1958.

[5] Frederick W. Taylor. The principles of scientific management. 1914.

[6] Thomas H. Davenport. Competing on analytics. *Harvard Business Review*, 84, 2006.

[7] Andrew McAfee and Erik Brynjolfsson. Big data: the management revolution. *Harvard Business Review*, 90, 2012.

[8] Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 1955.

[9] Herbert A. Simon. The new science of management decision. 1960.

[10] Irene Goll and Abdul M. A. Rasheed. Rational decision-making and firm performance: the moderating role of the environment. *Strategic Management Journal*, 18, 1997.

[11] J. W. Dean and M. P. Sharfman. Does decision process matter? a study of strategic decision-making effectiveness. *Academy of Management Journal*, 39, 1996.

[12] Erik Brynjolfsson, Lorin M. Hitt, and Heekyung Hellen Kim. Strength in numbers: How does data-driven decisionmaking affect firm performance? 2011.

[13] Gerrit Lahrmann, Frederik Marx, Tobias Mettler, Robert Winter, and Felix Wortmann. Inductive design of maturity models: Applying the Rasch algorithm for design science research. *Lecture Notes in Computer Science*, 6629, 2011.

[14] Erol Biberoglu and Hisham Haddad. A survey of industrial experiences with CMM and the teaching of CMM practices. *Journal of Computing Sciences in Colleges*, 18, 2002.

[15] James Bach. The immaturity of the CMM. *American Programmer*, 7, 1994.

[16] Tobias Mettler and Peter Rohner. Situational maturity models as instrumental artifacts for organizational design. *Proceedings of the 4th international conference on design science research in information systems and technology*, 2009.

[17] Jeffrey Pfeffer and Robert I. Sutton. Knowing "what" to do is not enough: Turning knowledge into action. *California Management Review*, 42, 1999.

[18] Jörg Becker, Ralf Knackstedt, and Jens Pöppelbuß. Developing maturity models for IT management. *Business & Information Systems Engineering*, 1, 2009.

[19] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *Management Information Systems Quarterly*, 28, 2004.

[20] Barbara Kitchenham. Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33, 2004.

[21] John C. Ory. Meta-assessment: Evaluating assessment activities. *Research in Higher Education*, 33, 1992.

[22] Daniel L. Stufflebeam and George F. Madaus. The standards for evaluation of educational programs, projects, and materials. *Evaluation in Education and Human Services*, 6, 1981.

[23] Daniel L. Stufflebeam. A note on the purposes, development, and applicability of the Joint Committee evaluation standards. *American Journal of Evaluation*, 25, 2004.

[24] Isabel E. Allen and Christopher A. Seaman. Likert scales and data analyses. *Quality progress*, 40, 2007.

[25] Watts S. Humphrey. Characterizing the software process: a maturity framework. *IEEE software*, 5, 1988.

[26] Wayne W. Eckerson. Gauge your data warehouse maturity. *Information management*, 14, 2004.

[27] Wayne W. Eckerson. TDWI's business intelligence maturity model. *The Data Warehousing Institute*, 2009.

[28] Tobias Mettler. Maturity assessment models: a design science research approach. *International Journal of Society Systems Science*, 3, 2011.

[29] Fern Halper and David Stodder. TDWI analytics maturity model guide. *The Data Warehousing Institute*, 2014.

[30] Bill Hostmann and John Hagerty. IT-score overview for business intelligence and performance management, 2010.

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

[32] Claudia Imhoff and Colin White. Self-service business intelligence. *Empowering Users to Generate Insights, TDWI Best practices report*, 2011.

[33] Bill Schmarzo. *Big Data MBA: Driving Business Strategies with Data Science*. 2016.

[34] Booz Allen Hamilton. The field guide to data science, 2015.

[35] Chris Nott. Big data & analytics maturity model — IBM big data & analytics hub. `http://www.ibmbigdatahub.com/blog/big-data-analytics-maturity-model`, 2014.

[36] Scott Rigby and Gerardo Contreras. Adobe analytics operational readiness playbook, 2015.

[37] Thomas H. Davenport and Jeanne G. Harris. Competing on analytics: The new science of winning. 2017.

[38] Browser Market Share. Browser market share. `https://netmarketshare.com`, 2018.

[39] Drift. Conversational marketing platform. `https://www.drift.com/`, 2018.

[40] Arent van 't Spijker. *The New Oil.* 2014.

[41] Yuri Demchenko, Paola Grosso, Cees de Laat, and Peter Membrey. Addressing big data issues in scientific data infrastructure. *Collaboration Technologies and Systems*, 2013.

[42] Alexandros Labrinidis and Hosagrahar V. Jagadish. Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5, 2012.

[43] Surajit Chaudhuri and Umeshwar Dayal. An overview of data warehousing and OLAP technology. *ACM Sigmod record*, 26, 1997.

[44] Vijay Khatri and Carol V. Brown. Designing data governance. *Communications of the ACM*, 53, 2010.

[45] Jan P. Albrecht. How the GDPR will change the world. *Eur. Data Prot. L. Rev.*, 2, 2016.

[46] Andy Neely, Mike Gregory, and Ken Platts. Performance measurement system design. *International Journal of Operations & Production Management*, Volume 15, 1995.

[47] Foster Provost and Tom Fawcett. Data science and its relationship to big data and data-driven decision-making. *Big Data*, 1, 2013.

[48] Steven Miller and Debbie Hughes. The quant crunch: How the demand for data science skills is disrupting the job market. 2017.

[49] Jake R. Carlson, Michael Fosmire, Chris Miller, and Megan R. Sapp Nelson. Determining data information literacy needs: A study of students and research faculty. *Libraries Faculty and Staff Scholarship and Research.*, 23, 2011.

[50] Dhruba Borthakur. HDFS architecture guide. *Hadoop Apache Project*, 53, 2008.

[51] Jeffrey Dean and Sanjay Ghemawat. MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 51, 2008.

[52] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, 11, 2009.

[53] Gary L. Neilson, Karla L. Martin, and Elizabeth Powers. The secrets to successful strategy execution. *Harvard Business Review*, 86, 2008.

[54] Power BI. Interactive data visualization BI tools. `https://powerbi.microsoft.com`, 2018.

[55] Bradley R. Staats, David J. Brunner, and David M. Upton. Lean principles, learning, and knowledge work: Evidence from a software services provider. *Journal of operations management*, 29, 2011.

[56] James Bennett and Stan Lanning. The Netflix prize. *Proceedings of KDD cup and workshop*, 2007.

[57] Modern Analytics. Predictive analytics solutions & automated big data. `https://www.modernanalytics.com/`, 2018.

[58] Deloitte. Amazon-whole foods acquisition: Six competitive realities for retailers. https://www2.deloitte.com/us/en/pages/consumer-business/articles/amazon-whole-foods-acquisition.html, 2018.

[59] Michael J. Franklin, Sailesh Krishnamurthy, Neil Conway, Alan Li, Alex Russakovsky, and Neil Thombre. Continuous analytics: Rethinking query processing in a network-effect world. *CIDR*, 2009.

[60] Jit Kee Chin, Mikael Hagstroem, Ari Libarikian, and Khaled Rifai. Advanced analytics: Nine insights from the C-suite — McKinsey & Company. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/advanced-analytics-nine-insights-from-the-c-suite, 2018.

Chapter 11

Appendices

11.1 Example Report

The report starts on next page.

Data-Driven Maturity Report

Score: 2.8/5

Organization:	Acme Inc.
Evaluator:	John Doe
Role:	Chief Information Officer
Date:	22nd of May, 2018

Data: 3.7

Metrics: 4.3

Skills: 1.7

Technology: 2.0

Leadership: 2.5

Culture: 2.0

Strategy: 3.0

Agility: 2.0

Integration: 4.0

Empowerment: 3.0

Reading Guide

This report contains the results of the data-driven maturity assessment. The first part contains a brief visual summary of the main scores that can be used to quickly communicate results. The other part consists of a personalized guide to data-driven maturity. The maturity score for each dimension is used to select the most relevant information out of the full data-driven maturity model. Some caution has to be applied in order to correctly interpret these results. The purpose of this assessment is to gain insight into where your organization is standing, and not to compare your results to other organizations. This is because the view of the assessor influences the results. The assessment should also not be used as a measurement of performance. Using this assessment as a measurement of performance would encourage an optimistic bias in filling out the results. No definitive conclusions should be made based on the results of this assessment alone. This assessment serves as a starting point for discussion, not a conclusion. The assessment may not fully take into account all contextual factors and nuances.

Data Optimizing → Empowering 3.7

Data is already being collected on a massive scale. You probably collect everything there is to collect about your customers. Keep looking out for new opportunities to learn more. Third-party data is another opportunity to get access to valuable data. External companies, specialized in data collection, can sell you market data that complements your own data. Third-party data is often expensive, so make sure it is going to be worthwhile. Data governance is becoming very important if you are starting to grant users, across the organization, access to data. You want employees to have access to all the relevant data, but you also have security concerns. With many access points and a large data volume a clear governance strategy is necessary.

- Look for new data opportunities, including from third-party sources
- Create a clear data governance strategy that scales

Metrics Empowering → Innovating 4.3

Metrics are being collected everywhere in the organization, and are being actively used to improve products and processes. KPI's are standardized across the organization. Advanced analytical KPI's are used for strategic goals, and are leading indicators for company performance. They allow organizations to better focus their efforts and measure the success. New KPI's may be introduced that are more in line with the redefined business model of the organization. KPI's have the ability to clearly communicate and incentivize priorities within the organization. Metrics are also being used more broadly to track internal processes and goals, such as employee performance or satisfaction. Metrics found from methods such as cohort analysis can be useful to measure the progress of new ventures. Innovation accounting is a technique to measure the success of new ventures instead of relying on traditional accounting methods, which may not be suitable for measuring the success of new innovation.

- Standardize KPI's across the organization
- Track new ventures using innovation accounting methods, instead of traditional metrics

Skills Reporting → Analyzing 1.7

Analytics evolves from showing numbers to calculating numbers, and the required skills evolve with it. Data mining and analytical skills are the most important skills in this phase. The current team has to be trained for more advanced analysis and the creation of analytical models that can be used for predicting future events based on new incoming data. In addition to basic statistics, advanced knowledge of data and data mining techniques is necessary. The retention of current talent is also a priority, because it can take a lot of resources to bring newly hired employees up to speed. A small team in the beginning means that a few individuals will have critical knowledge about the systems. Expanding the team will spread the critical knowledge and help overcome a sudden departure.

- Train or attract data scientist to dive deeper into the data.
- Retain talent with critical knowledge.
- Introduce employees to basic data literacy skills.

Technology Analyzing → Optimizing 2.0

The technological challenge is to integrate these advanced machine learning algorithms in production systems. There is a massive amount of data and calculation necessary to train these models. Timing is a big challenge: New data sometimes has to lead to insight only moments later. The infrastructure has to be able to quickly transport data. Batch processing has to make way for streaming data. From a pull structure, i.e. the data is queried when needed, to a push structure: When new data is generated, an event, it is pushed and processed immediately. Event-driven architectures are build to react to and handle incoming events. Personalization is a big part of optimizing the customer experience. Automated integration of analytics allows us to adjust the customer journey for every unique individual. This is unfeasible to do by hand for everyone. Machine learning algorithms can prescribe the best journey for every customer. Data collection about the customer is needed to serve as input to the analytical model.

- Invest in technology that can automatically process new emerging data in near real-time.
- Invest in technology to automatically integrate analytical insight into production systems.

Leadership Analyzing → Optimizing 2.5

Leadership has the important task to lead the organization across the chasm from 'doing analytics' to 'breathing analytics'. Great progress has already been made, and the organization might feel it is no longer behind the pack. This may lead to a reduced sense of urgency and less commitment to invest and expand analytical activities. But, we do not want to use analytics to do the same thing, we want to use analytics to make a difference. To make a difference and compete on analytics, you need to continually improve. So this is a critical step for the organization: Are you content with the current situation, or do you go above and beyond?

- Bridge the chasm from doing analytics as a side activity to being a cornerstone of all processes.
- Encourage the organizational culture to be the driving force for analytics, instead of leadership.

Culture Analyzing → Optimizing 2.0

Culture is probably one of the hardest things to change in any organization. Changing culture is a grind, and there is no magic potion. A culture that embraces analytics is necessary to sustain analytical activities. Leadership provides the runway for take-off, but Culture keeps it flying. At this stage a culture of continuous improvement, or kaizen, should be stimulated. A culture where everyone is engaged to improve the organization goes hand in hand with our philosophy for an analytical organization. Analytics is born out of the need to break the status quo, and to look for opportunities to improve. Analytics can provide the tools to the organization to analyze the current situation, and improve upon that. Analytics can support a culture of continuous improvement, and that culture can in turn sustain an analytical culture.

- Stimulate a culture of continuous improvement through data.
- Encourage experimentation to test new optimizations.

Strategy Optimizing → Empowering 3.0

The strategy of the organization should strive for the adoption of analytics in all business processes, and new products and services. Data and analytical activities form a central part of the strategy. Data is recognized as a valuable asset that can used or potentially sold. Analytics is used to optimize existing processes, products, and services. New products and services are augmented with 'smart' capabilities that add value for the customer. A strong analytical culture is a major strength that competitors might not have. Continuous improvement and innovation through analytics allow the organization to be competitive. A comprehensive customer view provides a way to better fit products and services to the customer.

- Strive to adopt analytics in every process and product
- Compete with strong analytics and continuous improvement

Agility Analyzing → Optimizing 2.0

Analytical processes are getting more interwoven with business processes. Strong execution in the form of collaboration and communication are necessary to create a fluid process. Clear decision and responsibility boundaries should be in place. Continuously rethink if the right people have the responsibility over a certain domain. Agile organizations are able to change fast in part due to a culture that embraces change and innovation. Introducing analytics changes the a business process, so it could be beneficial to change the owner of the process.

- Define or restructure clear decision and responsibility boundaries.
- Embrace change in order to innovate and react faster.

Integration Empowering → Innovating 4.0

We see analytical results further integrated into all business activities. It tells us what to do on a macro-scale: what business ventures to pursue, and on a micro-scale: what do we recommend this customer?. Analytics moves further to the point where data is generated. For example an IoT devices can already have machine learning capabilities built-in. The advantage is that the data is generated and analyzed on the same device. The computational efforts are also distributed. This architecture allows for massive scaling.

- Use analytics to make all decisions on a strategic scale and on an individual customer scale
- Integrate analytical processing capabilities into all products without centralizing data collection

11.2 Assessment Application Homepage

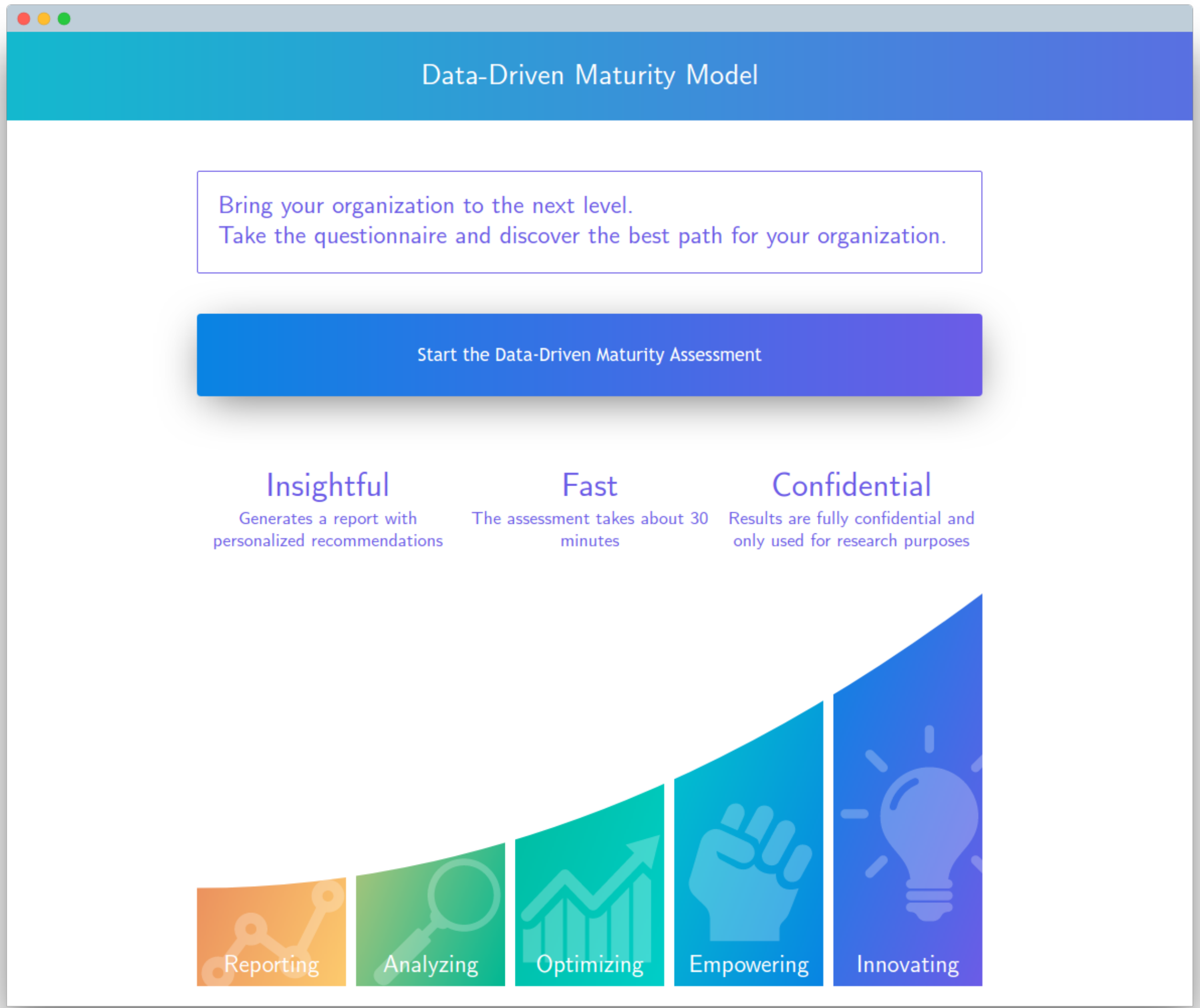

Figure 11.1: Homepage of the assessment application.