Universiteit Leiden

ICT in Business

Measuring the status quo of big data analytics in Dutch academic hospitals

Name: Simone Cammel
Student-no: 1039725

Date: 28/06/2017

1ST SUPERVISOR: ARNO KNOBBE

2ND SUPERVISOR: TINO DE RIJK

ABSTRACT

Big data analytics is expected to have great potential for the healthcare industry. Some applications are precision medicine, disease prevention and personalized care. However, there are still challenges to overcome that are specific for the Dutch healthcare industry (Ottenheijm, 2015). Nictiz identifies challenges on technology, standardization, data access and privacy concerns. These challenges might also be faced in other countries but were specifically identified for the Dutch industry.

This paper presents a maturity model to assess the big data maturity of a Dutch academic hospital that specifically addresses the challenges for the industry. Academic hospitals in the Netherlands consist of both medical specialists and researchers. Researcher are focussed on the long-term as they discover new insights in the medical sector that patients will benefit from years from now. On the contrary, there are medical specialists whose focus lies more in short-term solutions as they want the patient to be cured as soon as possible.

The maturity model has seven domains, Strategic Alignment, Data Governance, Information Technology, Data, Organization, Privacy and Innovation. The model provides the current maturity of a Dutch academic hospital and the to-be maturity in two years. This model was tested and validated at three different academic hospitals in the Netherlands. The current maturity of the assessed hospitals is between 2,6 and 2,9 and the to-be maturity of the assessed hospitals is between 3,0 and 3,6 on a scale of 1 to 5.

It will take more than two years to reach a higher level of maturity as change in these massive organizations takes time. The highest level of maturity will most likely not be reached anytime soon because the academic hospitals do not use the same standards or infrastructure. Terms as data lakes or distributed data storages are not even thought of by the assessed academic hospitals. The added value of big data analytics is not considered by most employees in these academic hospitals. There is a preference for statistical approaches over explorative analysis.

The Dutch academic hospitals are now collaborating with other academic hospitals and participating in many different national initiatives that contribute to preconditions for a high maturity of big data analytics in healthcare. This leads to many different standards, infrastructures and parties working on enabling healthcare data sharing. The national initiatives and the academic hospitals consider healthcare data and research data as two separate worlds. As a result, opportunities for both these disciplines can be missed.

Keywords: Big data analytics; maturity model; healthcare; academic hospitals;

CONTENTS

1. INTRODUCTION

Big data is a term that is often used as a *buzzword***.** The term has become ubiquitous as academia, the industry and the media all use other definitions of the term big data. The expectation is that the future for healthcare lies in leveraging data (Shaffer & Craft, 2016). Precision medicine, a technique based on big data analytics, is mentioned as healthcare's ultimate manifestation of digital business. Precision medicine is an approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person.

To be able to understand the possibilities of big data for the healthcare industry, an overview of the existing definitions of big data analytics is first introduced. A definition for big data analytics is derived from these existing definitions, which will be used in this research project. Applications of big data analytics in the healthcare industry are described and the challenges to adopt big data analytics in healthcare are discussed. This chapter concludes with the scope for this research project and the research questions for this project. Finally, the chosen research approach and thesis outline are described.

1.1BIG DATA ANALYTICS IN DUCTH ACADEMIC HOSPITALS

Big data is defined by the U.S. congress in August 2012 as "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information". Big data analytics are recognized in the Netherlands as an opportunity for healthcare (Koumpouros, 2014). Vast amounts of data are generated in healthcare today. This data has the potential to improve the healthcare industry. Nictiz, the Dutch competence center for standardization and eHealth, recognized these opportunities (Ottenheijm, 2015):

- Opportunities for research such as precision medicine
- Disease prevention by tracking patients behavior with i.e. activity trackers
- Qualitative care by predicting which treatment will work best
- Personalized care by using unique characteristics of an individual to determine a specific care plan
- Public care, for example by predicting the next outbreak of a spreading virus
- Fraud detection

However, they also define challenges that need to be overcome for big data analytics to reach its full potential in the Dutch healthcare system. These challenges are defined as:

Technology
The traditional ICT architecture currently in use has a lot of limitations when it comes to big data. Some elements that are currently missing are the ability to store massive amounts of data, analysing it and visualizing the outcome. Various data assets need to be combined for a thorough big data analysis which means many parties have to share date with one another. Agreements need to be made on what technologies will be used.

Standardization
Currently there is no standardization between healthcare providers, previously researched for a competition (de Rijk, 2016) There are many different systems using different formats. Standardizing these formats and defining data definitions would be a step in the right way.

Data Access
A very important challenge to overcome is the fact that data is currently stored in a fragmented way.

Every department within a hospital, every healthcare provider and even each patient could have its data stored in a silo, while big data initiatives can only succeed when data is combined. Now that patients are more and more aware of the fact that they own their data, it is still a long way to go to change the mind-set towards a shared data-platform.

Privacy

Another very important discussion about applying big data to healthcare is ownership of information and who can use this information and for what purposes. Since the Electronic Patient File system[1] (EPD) is in place, there has been a lot of resistance towards the project. Doctors think patients are wary of sharing their data and breaches of their trust can be disastrous for big data initiatives.

The Dutch academic hospitals have three tasks they need to fulfil: develop knowledge by doing ground-breaking research, apply this knowledge in the most complex healthcare cases and spreading this knowledge via education to (future) healthcare professionals (Over de UMC's, 2017). The NFU, the Dutch federation of University hospitals represents the eight cooperating UMCs in the Netherlands, arranged a meeting on big data on June 16, 2016. Now, there is an initiative from the NFU, called *Data4lifesciences,* which aims to make clinical and experimental data on all UMC-related patient available to others.

Another aspect of privacy that is specific for the healthcare industry is that data that is used for medical research has to be pseudonymized. The data used for the analysis should be anonymous, but in case of important findings individuals should be notified. Thus, it should also be possible for data to be de-anonymized which is different from other industries where data only has to be anonymized and not pseudonymized.

1.2 MOTIVATION

The previous paragraph described the potential of big data analytics for the Dutch healthcare industry and the challenges that will need to be overcome for big data analytics to reach its full potential in the Dutch healthcare system. There are many national initiatives to overcome these challenges, but they are not all supported by every UMC. The result is a scattered landscape with different parties trying to work separately on solving one of these challenges. Meanwhile, UMCs are also developing their own infrastructure to support big data analytics for their own researchers. The current situation is discussed in Chapter 6.

Scientific contribution

Whereas big data analytics in healthcare is widely covered in scientific literature (an initial search in google scholar results in around 500.000 hits), maturity models on big data in healthcare (or a variation on this) result in less than 1000 hits. There is no scientific method to assess the maturity of big data analytics in healthcare, and this has never been scoped to the Dutch academic hospitals. The found maturity models are discussed in Chapter 4.

Practical contribution

Big data analytics in healthcare is a promising topic that should help hospitals to provide the right care at the right price. The healthcare society has lagged behind in the use of big data compared to other industries (Raghupathi & V, 2014), while there are major forces in the healthcare industry such as a huge cost pressure from the government and a rapidly increasing number of patients that are in need of medical care each year.

[1] Elektronic Patient Filing system https://nl.wikipedia.org/wiki/Elektronisch_pati%C3%ABntendossier

As there is not one clear defined path how to take an academic hospital to a next level on this matter, a maturity model is proposed to provide guidance on this matter. This maturity model is described in Chapter 5.

Finally, the status quo of the UMCs is compared, the status quo of the national initiatives is compared and finally these two are compared to see if the strategies from current national initiatives on big data analytics and the academic hospitals are aligned. An advice is provided to the Dutch academic hospitals on aligning the strategies and working together between academic hospitals and national initiatives towards a high maturity of big data analytics in (and between) the academic hospitals in Chapter 7.

1.3 SCOPE

The scope of this project are big data analytics in Dutch academic hospitals and the national initiatives they are involved in to organize big data analytics on data from all the UMCs. The healthcare industry is enormous, with more than 1.2 million people working in healthcare in the Netherlands. Academic hospitals in the Netherlands have the responsibility to work on innovative ground-breaking research to improve healthcare. Thus the expectation is that the Dutch academic hospitals are most mature on big data analytics. The maturity of three academic hospitals was measured. These hospitals were the VUMC in Amsterdam, the Radboud UMC in Nijmegen and the LUMC in Leiden. The other 5 academic hospital in the Netherlands were not assessed because not all UMCs wanted to cooperate and due to time constraints. Peripheral hospitals and academic hospitals in other countries were out of scope.

Six interviews with representatives from different initiatives were held to get an overview of the national initiatives to support big data analytics in healthcare. Due to time constraints it was not feasibily to cover all these initiatives with interviews.

1.4 RESEARCH QUESTIONS

In the previous subsections the challenges that need to be solved for big data analytics to succeed in the Dutch healthcare industry were briefly described. The aim of this research project is to answer the question: ‘What is the status quo of academic hospitals on big data analytics?’

The current state of the academic hospitals on big data analytics will be assessed. Maturity models are designed for this purpose: to provide a means to assess and rate the maturity of a selected domain. Also, the national initiatives that try to contribute to big data analytics in healthcare are assessed and finally suggestions are made for improvements of the maturity and to align these both worlds.

The research project is split up in four parts. First, the topic big data analytics specifically for the Dutch academic hospitals is explored and characteristics of this industry are provided. Secondly current maturity models are evaluated on their ability to capture the characteristics of the Dutch academic hospitals in the model. Thirdly, the maturity of the academic hospitals in the Netherlands is assessed on their maturity and finally this is compared to the national initiatives on big data analytics in healthcare and how these worlds relate to each other. The research questions are:

Research question 1	What are characteristics of big data analytics for the Dutch academic hospitals that a maturity model should capture?
Research question 2	Is there a big data maturity model that meets the characteristics specific for the Dutch healthcare industry?

Research question 3	How mature is big data analytics in Dutch academic hospitals currently?
Research question 4	How do the Dutch academic hospitals and the national initiatives on big data analytics in healthcare relate and how can they reinforce each other?

The first question is answered in Chapter 3, the second in Chapter 4, the third in Chapter 5 and the fourth in Chapter 6.

1.5 RESEARCH APPROACH AND METHODS

The adopted research approach was a combination of qualitative and quantitative design science approach. The first design of the maturity model was validated by an expert group that are not employed by an academic hospital and two case studies at academic hospitals in the Netherlands. The model was updated with the feedback from these three sessions. The revised version of the model was tested in another case study at an academic hospital. The research approach and method is discussed in more detail in Chapter 2.

1.6 THESIS OUTLINE

The structure of this thesis is as follows. Chapter 2 discusses the research approach and method. Chapter 3 provides a literature overview of (definitions on) big data, big data healthcare and specifically big data in the Dutch academic hospitals. Chapter 4 consists of a critical analysis of existing maturity models on this topic. In Chapter 5, the developed maturity model is discussed. The results of the case studies are discussed in Chapter 6, as well as an overview of the current national initiatives and their vision on the future. Finally, the research project is concluded and suggestions for future research are provided in Chapter 7.

2. METHODS

In this chapter the methodology used to answer the research questions is explained. First the chosen approach, design science, is discussed. Then a general overview is given with all steps in the research process and where the results of these steps can be found. Each step is elaborated on in more depth.

2.1 DESIGN SCIENCE

A design science approach with case studies was used. Design science research is usually focused on explaining and improving the current situation and often used when developing maturity models (Wendler, 2012). The key aspect of design science is to develop and to evaluate artifacts. In this case the artifact is the maturity model. The process model for maturity model development as proposed by Becker et al was used to design the artifact (Becker, Knackestedt, & Pöppelbuss, 2009).

Design science is iterative as improving situations tends to need many iterations. This research project consisted of two iterations of the design cycle which will be discussed in the next sections.

2.2 OVERVIEW OF THE RESEARCH APPROACH

Figure 1 provides an overview of the adopted research approach. Interviews and case studies are qualitative in nature, but these were enriched with quantitative data. The purpose of each of the research approaches is discussed in this section.

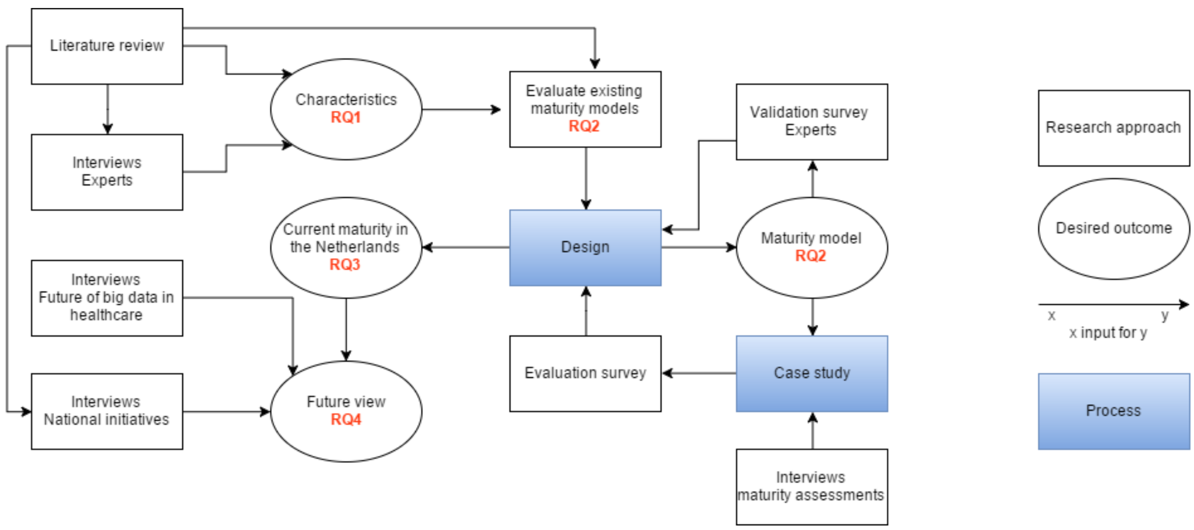


FIGURE 1 OVERVIEW RESEARCH APPROACH

The purpose of a literature review is to get an idea of the topic and learn about relevant studies previously done on this topic. The literature review is mainly done for the following reasons:

- To get familiar with the topic *big data analytics in Dutch academic hospitals* and its context, including characteristics specific for this industry that can be used to measure the big data maturity of these academic hospitals
- To find maturity models on this topic that could be used to assess the maturity of the academic hospitals or used as an inspiration for the design of such a model
- To find national initiatives on big data analytics in healthcare

The context found in the literature review is used as input for the interviews with experts on the topic. Interviews allow for in-depth information. The purpose of these interviews was to get an in-depth view on the context and to find more characteristics specific for this industry.

The characteristics and the found maturity models were used as input for the maturity model evaluation phase. All found maturity models were evaluated using the specified characteristics.

The evaluation of the maturity models is used as input for the design. The design was then evaluated to measure how well the artifacts support the problem. This evaluation was done in two different ways: asking experts and using case studies.

The experts were asked to evaluate the maturity model in a survey. The survey was mainly done for the following reasons:

- To find out if the maturity levels were relevant, mutually exclusive, complete and accurate
- To find out what domains were relevant, mutual exclusive, complete and accurate
- To find out what attributes were relevant, mutual exclusive, complete and accurate
- To find out whether the maturity model was understandable
- To find out whether the maturity model was useful and practical
- To get an outside-in perspective on the maturity model

Case studies were done at three academic hospitals in the Netherlands. In the first iteration two case studies were conducted simultaneously and in the second iteration one case study was conducted. A case study consisted of several interviews with different employees in one academic hospital per case study. The purpose of these interviews was to be able to get answers to the different questions of the maturity model to determine the maturity of the academic hospital.

The outcome of these interviews was then used to determine the maturity of the academic hospital. The result was given back to all interviewees, with a URL to an online survey to evaluate the outcome of the assessment. The survey was mainly done for the following reasons:

- To validate the outcome of the assessment
- To find out if the maturity levels were relevant, mutual exclusive, complete and accurate
- To find out what domains were relevant, mutual exclusive, complete and accurate
- To find out what attributes were relevant, mutual exclusive, complete and accurate
- To find out whether the maturity model was understandable
- To find out whether the maturity model was useful and practical
- To get an insider perspective on the maturity model

The results of the expert survey and the validation surveys were used to redesign the model. Another case study was then performed at another academic hospital, with another online survey to evaluate the outcomes of the assessment. The survey was then used to finalize the maturity model.

Finally, all maturity assessments were used as input for the overview of the current maturity in the Netherlands. A sounding board group was formed with experts on big data analytics in healthcare to define the future of big data analytics in healthcare.

Interviews with representatives from national initiatives on big data analytics in healthcare were interviewed. The purpose of these interviews was

- To find out what national initiatives there are
- To find out what the relation between initiatives is
- To find out what their strategy is for the years to come

The overview of the current maturity and the interviews on the future of big data analytics in healthcare and the interviews with representatives from national initiatives on big data analytics were used to define the future of big data analytics for Dutch academic hospitals.

In Table 1, each research question is linked to a research method

Research Question	Method	Discussion results
RQ1: What are characteristics of big data analytics for the Dutch academic hospitals?	Literature review, Interviews with experts	Chapter 3
RQ2: Is there a big data maturity model that meets the characteristics specific for the Dutch healthcare industry?	Literature review, Interviews with employees of UMCs, Expert validation survey, Case study evaluation surveys	Chapter 4 and 5
RQ3: How mature is big data analytics in Dutch academic hospitals currently?	Case study evaluation surveys	Chapter 6
RQ4: How do the Dutch academic hospitals and the national initiatives on big data analytics in healthcare relate on big data analytics and how can they reinforce each other?	Interviews on future of big data analytics in healthcare, Interviews with representatives from national initiatives	Chapter 6 and 7

TABLE 1 LINK BETWEEN RESEARCH QUESTIONS AND RESEARCH METHODS

In the section below, each method is discussed in more detail.

2.3 LITERATURE REVIEW

The literature review consisted of different types of literature to be reviewed. There were three different types of literature needed for this research: literature on big data analytics in healthcare, literature on big data maturity models and thirdly national initiatives to support big data analytics in healthcare. The results of the literature review are shown in Chapter 3.

2.3.1. LITERATURE ON BIG DATA ANALYTICS IN HEALTHCARE

Firstly, literature on big data was explored to understand the topic and current challenges in applying big data analytics. Then the search was narrowed to big data analytics in the healthcare industry. Both academic literature and 'grey' literature, which is not always peer-reviewed, were included. The grey literature was added to overcome the gap between research and practice. This is also in line with the adopted design science research approach.

Searches were done in Google Scholar for papers and books on *big data analytics, big data, advanced analytics, big data warehousing,* and *big data in organizations*. This resulted in mostly academic literature, but also grey literature. These papers were sorted on relevance and selected on the match of keywords in the title of these papers and their abstract.

In the second phase, the search was deepened by searching for big data analytics specifically in healthcare. Again, Google Scholar was used to find academic literature. Used queries spanned *big data analytics in healthcare, business intelligence in healthcare, big data medicine, big data hospitals, big data gezondheidszorg,* and *data analytics healthcare*. These papers were sorted on relevance and selected on the match of keywords in the title of these papers and their abstract.

Google Search and the Garner online (payed) database were used to find grey literature. The same queries as previously mentioned were used. The found articles were sorted on data to get a view on the current ideas on the topic by research organizations such as IBM, McKinsey, Gartner, HIMMS Analytics.

Finally, Google Search and Google Scholar were searched on big data analytics in healthcare in the Netherlands and big data analytics in academic hospitals in the Netherlands to get an idea of the current situation on the matter. The search on Google scholar had 7.240 results. The papers were judged on their title: is this truly about big data in healthcare in the Netherlands? If so, the abstract was read and judged on the potential value of the paper: will this add information to the previously found papers? The report from Nictiz (Ottenheijm, 2015) was found.

2.3.2. LITERATURE ON BIG DATA MATURITY MODELS

A quick search on maturity models learned that many different maturity models exist. Two different sorts of literature on maturity models were searched for. Firstly, literature on the process of designing a good quality maturity model was searched. Papers were found on Google Scholar for books and papers with the following queries: *understanding maturity models, developing maturity model, structure of maturity model, useful maturity model, designing maturity model, phases of maturity model*, *testing maturity models, questionnaire maturity models, evaluating maturity models.*

The scientific literature found provided sufficient knowledge on maturity models. This knowledge was used to assess maturity models on big data in healthcare.

A first search on big data analytics maturity models for hospitals resulted in only one academic paper, so a top-down approach was adopted. This top-down approach was adopted by first searching for general maturity models, and then narrowing the scope of the models in steps. Many maturity models were found, of which most are based on the CMMI. The CMMI is a popular maturity model that is focussed on optimizing operational processes developed by Carnegie Mellon University. In the next phase, maturity models specifically for big data, business intelligence and data warehousing were searched. Finally, big data analytics maturity models and business intelligence maturity models were searched specific for healthcare. Models that were selected for an extensive assessment were chosen by looking at titles, abstract, number of references made to the model and extensiveness of the description of the model to make a short-list. These models were assessed in their quality and if they could function as a foundation for our maturity model. The selection of models is discussed in Chapter 4.

2.3.3. LITERATURE ON NATIONAL INITIATIVES SUPPORTING BIG DATA ANALYTICS IN HEALTHCARE

In the starting phase of this research project, a conversation with IT architects from the LUMC made it clear that there are many national initiatives supporting big data analytics in healthcare. This needed a more direct approach to find literature on these initiatives, as a query on Google search on *koepels zorg data* does not give any results. The CIO from the LUMC and an IT architect from the LUMC provided a list with all national initiatives on this matter. Each of these national initiatives was found using Google Search. The national initiatives will be discussed in Chapter 6.

2.4 INTERVIEWS

Several interviews were conducted during this research project, with different purposes. These different types of interviews were:

- Interviews with experts in the field of big data in healthcare
  An expert in this regard is a consultant with an outside-in perspective on big data in healthcare. Attempts were made to answer all questions shown in Appendix C:
- Interviews with representatives from national initiatives on (big) data in healthcare
  A representative was a board member of the national initiative or an IT architect of the national initiative. Attempts were made to answer all questions shown in Appendix C.

- Interviews with UMC employees as part of the case study
  Multiple employees dealing with one (or more) of the domains from the maturity model were interviewed per case study. With the interviewees answers the questions from the maturity model were answered. These can be found in Appendix B.

Beforehand, each interviewee is asked for permission to record the interview. The recording was used to transcribe the interview to get an accurate transcription of the interview.

There were also interviews with a visionary team. These visionaries were experts on big data in academic hospitals with an inside perspective: they were all working in academic hospitals in the Netherlands. They are considered visionary because they expressed ideas on the future of academic hospitals and healthcare in the Netherlands. Because they expressed these ideas, they were asked for interviews or group sessions specific on the future of healthcare. . Some of these interviews were one-to-one. There were three sounding board group sessions where one group of visionaries discussed the future of big data in healthcare. More information on the interviews and the results of the sounding board group session can be found in Appendix C.

A complete list with all interviewees can be found in Appendix C

2.5 SURVEYS

The evaluation and validation of the maturity model was realized with a survey. The expert group that was also interviewed was sent a survey to evaluate the maturity model constructs and instruments. The survey was based on an evaluation template for expert review of maturity models (Salah, Paige, & Carins, 2014). The experts were asked to evaluate the first version of the maturity model. The questions were on sufficiency, accuracy, relevance, comprehensiveness, mutual exclusion, understandability, ease of use, usefulness and practicality. The questions of the survey and the results can be found in Appendix D. The results are discussed in Chapter 6.

All interviewees from UMCs that were part of the case study were also sent a survey when the first version of the report on the case study was finished. This survey included the questions that were sent to the expert group. They also were asked their opinion on each chosen answer to each question of the model. These questions and the results can be found in Appendix C. The results are discussed in Chapter 6.

2.6 DESIGN

The design of the maturity model was done iteratively. All steps in the process are described below. An overview of the complete design approach is visible in Figure 2.

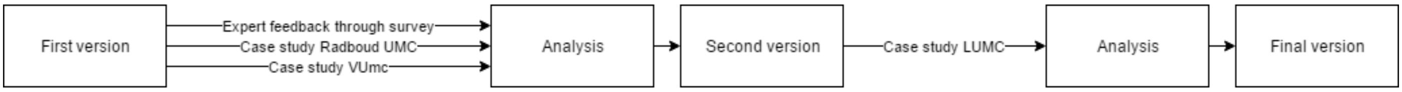


FIGURE 2 OVERVIEW OF THE DESIGN APPROACH

STEP 1: DESIGN THE FIRST VERSION

In the literature review no maturity model to measure big data analytics of academic hospitals in the Netherlands was found. Other models did not meet characteristics of the industry or other requirements. Based on the literature study and the expert interviews, a first version of the maturity model was designed. It is attached in Appendix B.

STEP 2: EXPERT FEEDBACK THROUGH SURVEY

The experts were asked for feedback on the maturity model through a survey as described in Chapter 2.5.

STEP 3: CASE STUDY AT RADBOUD UMC
The model was tested with a case study at the Radboud UMC. The structure of the case study is described in Chapter 2.7.

STEP 4: CASE STUDY AT VUMC
The model was tested with a case study at VUMC. The structure of the case study is described in Chapter 2.7.

The case studies were done simultaneously and not one by one because it was very time-consuming to perform one case study. The case study was time-consuming due to several reasons:

- Getting a sponsor in the academic hospital is hard to find, as these are very large organizations with 7000+ employees. It was difficult to find a good contact in each of the UMCs
- Scheduling interviews was difficult, as the interviewees have busy schedules. It was difficult to find an opening in their agenda
- It took weeks to get responses to the surveys due to a high workload from the interviewees

STEP 5: ANALYSIS OF FEEDBACK AND IMPROVEMENT
The results of the surveys and the feedback that was received over e-mail from both the experts and the two case studies were analyzed. These results were used to update the maturity model. The second version of the model is attached in Appendix B.

STEP 6: CASE STUDY AT LUMC
The new version of the model was testes with a case study at the LUMC. The structure of the case study is described in Chapter 2.7.

STEP 7: ANALYSIS OF FEEDBACK AND IMPROVEMENT
The results of the surveys and the feedback that was received over e-mail from the case study were analyzed. These results were used to update the maturity model. The final version of the model is attached in Appendix B. The results are discussed in Chapter 6.

2.7 CASE STUDIES

GATHERING INFORMATION
The maturity model was tested three times with case studies at three different academic hospitals in the Netherlands: Radboud UMC, VUMC and LUMC. Each case study was structured following the same process. First, a sponsor was appointed who would act as point of contact during the whole process. Together the appropriate persons to interview were decided. The list contained 6 persons at Radboud UMC, 3 at VUMC and 7 at LUMC. The low number of interviews at VUMC is partly due to a change of CIO during the case study. Interviews with all the interviewees were planned. Each interview lasted 45 to 90 minutes. When interviewees referred to documentation, these were sent by email after the interview.

ANALYSIS
All the information obtained by the interviews and the documentation were used to analyze the academic hospital. The questions of the maturity model were answered. The conclusions and recommendations were written down in a report.

EVALUATION
The report and the answers to the questions of the maturity model were evaluated with a survey that was sent to each of the interviewees. The survey is described in Chapter 2.5.

ADJUSTMENTS
The responses to the surveys were analyzed and used to adjust the maturity of the academic hospital. The report was updated. All the material was sent to the academic hospitals. The case study was ended with a presentation on the case study for all interviewees involved in the case study.

The results of the case studies are described in Chapter 6.

3. BIG DATA ANALYTICS IN ACADEMIC HOSPITALS IN THE NETHERLANDS

In this chapter, the results of the literature study and the interviews with experts on big data analytics will be discussed. The chapter is divided in several sections. In the first section, an overview of big data definitions in literature is provided. A definition of big data analytics for this thesis is presented. Secondly, literature on applications and limitations of big data analytics in healthcare are discussed and, thirdly, the literature results on big data analytics in Dutch academic hospitals are presented. The literature study serves to satisfy the following knowledge goal:

- To get familiar with the topic *big data analytics in Dutch academic hospitals* and it's context, including characteristics specific for this industry that can be used to measure the big data maturity of these academic hospitals

Finally, the interviews with experts on big data analytics in healthcare are discussed. The purpose of these interviews was to get an in-depth view on the context and to find more characteristics specific for this industry.

In the last section, we conclude with the characteristics found of big data analytics for the Dutch academic hospitals.

3.1 BIG DATA ANALYTICS

This section will discuss several definitions on big data and finally formulate one definition that will be used as a definition for this thesis.

"Every day 2.5 quintillion bytes of data are created – so much that 90% of the data in the world today has been created in the last two years alone" (IBM, sd). IBM claims this data as *big data.*

Big data
Big data is a term that has become quite ubiquitous (Ward & Barker, 2013). In academia, the industry and the media there are different definitions of big data analytics used. A definition from Gartner (Douglas, 2001), defining big data as the combination of three V's: volume, variety and velocity, is one of the most adopted definitions on big data. Others (IBM, 2015) add a fourth V: veracity.

- **Volume**
  Volume refers to the scale of data. Such an amount of data that can no longer be stored using *traditional* SQL databases but for example needs a distributed file system such as Hadoop.
- **Variety**
  Variety refers to the variety of data sources. There are two types of data: structured and unstructured data. Data is structured if it is categorized or placed in a logical structure. Twenty percent of all data is estimated to be structured, leaving eighty percent of all data unstructured (Grimes, 2008). An example of unstructured data is written text.
- **Velocity**
  Velocity refers to the speed of data creation. Data is now stored real-time, in massive amounts. Big data is constantly changing, as new data is stored constantly.
- **Veracity**
  Veracity refers to the quality of the data. This includes questions of trust and uncertainty with regard to data and the outcome of analysis of that data.

MiKE 2.0, the open source standard for information Management, defines Big Data by its size, comprising a large, complex and independent collection of data sets with the potential to interact (E & al., 2012). It cannot be handled with standard data management techniques due to the inconsistency and unpredictability of the possible combinations.

Big data analytics

In Figure 3 an overview of the conceptual architecture of big data analytics is shown. There are different sources, or a *variety* of data, that change frequently, a high *velocity* of data. Then there is a big data transformation from raw data to a transformed data format. This can be stored on a big data platform that is able to deal with a high *volume* of data. Finally, all this data can then be used for big data analytics using different applications.

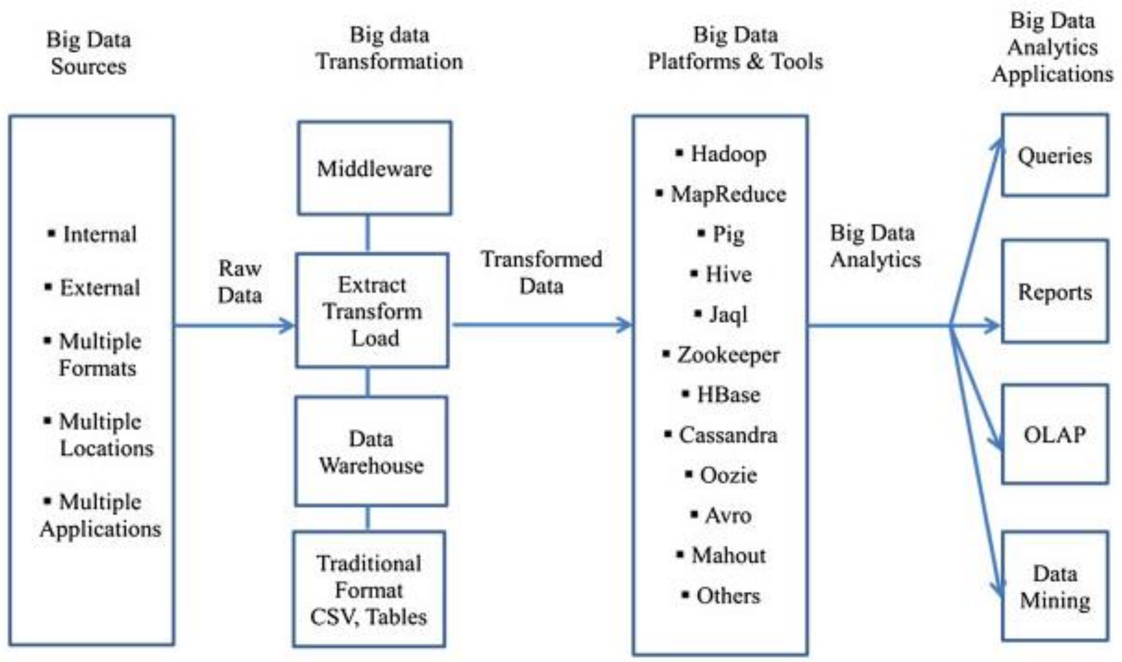


FIGURE 3 APPLIED CONCEPTUAL ARCHITECTURE OF BIG DATA ANALYTICS (GROVES, KAYYALI, KNOTT, & VAN KUIKEN, 2013)

Big data analytics is where advanced analytic techniques operate on big data sets (Russom, 2011). The term big data and analytics are often mentioned together. However, analytics can also be done on small data and this is often mistaken with big data analytics. Big data analytics are often mentioned with different terms such as advanced analytics or predictive analytics. A report from TDWI (Russom, 2011) shows that 43% of their survey respondents used a unique name for big data analytics, showing that there is no commonly accepted definition.

Data-driven vs Hypothesis-driven data analysis

Russom also describes a better term that could be used for big data analytics: discovery analytics or exploratory analytics. “With big data analytics, the user is typically trying to discover new business facts that no one in the enterprise knew before. To do that, the analyst needs large volumes of data with plenty of detail. This is often data that the enterprise has not yet tapped for analytics”. This is completely opposite from the traditional analytics methods, which are hypothesis-driven. With big data, a data-driven approach can lead to new insights and patterns that would not be thought of when forming a hypothesis (Shih & Chai, 2016). Big data can be used for hypothesis-driven approached, but ‘small’ data might also be sufficient for these cases. When a data-driven approach results in hypothesis, these can be validated using traditional data analytics methods.

For this research project we adopt the following definitions for big data and big data analytics:

Definition: Big data

Big data is defined as a high volume of data that can no longer be stored using traditional methods, has a high variety of data combining both structured and unstructured data and has a high velocity of data: the data set is constantly changing.

Definition: Big data analytics

Big data analytics is where analytical techniques are used to operate on big data with a data-driven approach for discovery analytics or exploratory analytics.

3.2 BIG DATA ANALYTICS IN HEALTHCARE

This section will discuss literature on applications and limitations of big data analytics in healthcare.

A changing industry

The healthcare industry is changing. People are having a longer life expectancy than 60 years ago, but they are also suffering longer from chronical illnesses (Suzman & Beard, 2011). Gartner releases a Hype Cycle every year, covering upcoming trends in a certain domain (Shaffer & Craft, 2016). The Hype Cycle for Healthcare providers from 2016 can be seen in Figure 4.

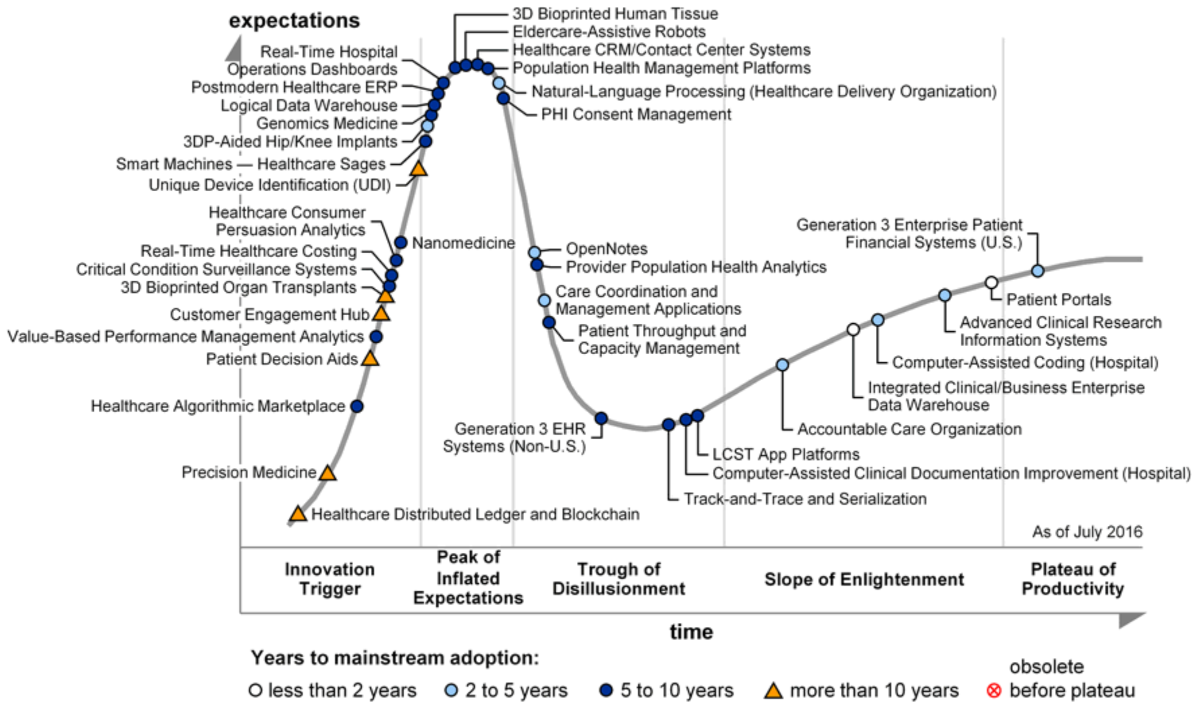


FIGURE 4 GARTNER HYPE CYCLE FOR HEALTHCARE PROVIDERS, 2016 (SHAFFER & CRAFT, 2016)

They conclude that the progress in healthcare lies in leveraging data. Precision medicine is only at the beginning of the hype cycle and not yet near the peak of inflated expectations. . Big data has been removed from the hype cycle while it was there until 2015, because it is now divided into specific uses of big data technologies in healthcare such as advanced clinical research information systems. Thus according to the Gartner hype cycle big data is ready for mainstream use, advanced clinical research information systems has reached the plateau of productivity, genomics medicine is reaching the peak of inflated expectations and precision medicine is still too new to be a hype.

Applications

There are many applications of big data analytics in healthcare described in literature. Shaffer & Craft mention that precision medicine is healthcare's ultimate manifestation of digital business. Precision medicine is an approach for disease diagnosis, treatment and prevention that takes into account individual variability of genes, physiology, environmental exposures and lifestyle (Shaffer & Craft, 2016). It combines advances in genomics medicine, wearables, electronic health record systems and mobile device applications. An example of precision medicine can be found at the end of this subsection.

Deloitte (Deloitte, 2015) predicts that the healthcare industry will transform to personalized health: treatments and medicine will not be made for the general public but customized for each individual. There will also be a shift from sick care to health care: prevention will become more important. The shift will change healthcare from being re-active to proactive. Medical specialists will no longer base diagnosis mostly on experience, but a shift will take place where diagnosing will be evidence-based.

Sun & Reddy agree with this opinion as they believe that decision making in healthcare will become evidence-based and big data analytics will make this change happen (Sun & Reddy, 2013).

That big data analytics will impact healthcare seems accepted in literature. However, expectations on where and how it is going to make its impact differs. McKinsey (Groves, Kayyali, Knott, & Van Kuiken, 2013) distinguishes five pathways that involve:

- **Right living**; give patients an active role into their own treatment and disease prevention. For example monitoring through personal devices
- **Right care**; give patients the most appropriate treatment. For example evidence-based medicine
- **Right value**; eliminate fraud, waste and abuse in the system
- **Right provider**; match skills to complexity of the task
- **Right innovation**; identify new therapies and opportunities

Nictiz, the Dutch national competence center for standardization and eHealth, recognizes slightly different categories of opportunities (Ottenheijm, 2015)

- Opportunities for research. For example making it easier to find patients for a trial
- Disease prevention
- Qualitative care
- Personalized care
- Public care
- Fraud detection

The terms that are used for the opportunities might vary, but there seems to be agreement on these potential applications of big data analytics in healthcare in literature (Feldman, Martin, & Skotnes, 2012), (Raghupathi & V, 2014), (Belle & Thiagarajan, 2015), (Bollier, 2010).

Example of precision medicine: Inflammatory bowel disease patient (Nguyen & Mendes, 2011)
Patients with the inflammatory bowel disease have considerable inter individual differences in efficacy and side effects of commonly used medications. Azathioprine is the first medicine that these patients receive, while 40% of patients fail to achieve a clinical remission on this drug and 15-28% experience adverse effects such as hematological and hepatic toxicities. Up to 29% of patients with severe ulcerative colitis fail to respond to steroid treatment and require surgery.

The clinical practice for ulcerative colitis is the following step-up approach:
- Aminosalicylate
- Prednisolone
- Azathioprine
- Mercaptopurine
- Leukapheresis
- Infliximab

This means that these drugs will be prescribed to the patient one by one, until a drug that is actually working is found.

Precision medicine could solve this ‘trial and error’ approach. There are polymorphisms that have been associated with increased risk. The genotype of a patient can already give insights in which of these drugs will work for a patient. The result would be that with this precision medicine approach, the patient will have the right drug at once instead of having to test drugs first until a working drug is found.

Challenges
There are still challenges that need to be overcome for big data analytics to become a standard practice in the healthcare industry. IBM (Sun & Reddy, 2013) addresses the following challenges for big data in healthcare:

- Inferring knowledge from complex heterogeneous patient sources. Leveraging the patient/data correlations in longitudinal records
- Understanding unstructured clinical notes in the right context
- Efficiently handling large volumes of medical imaging data and extracting potentially useful information and biomarkers
- Analyzing genomic data is a computationally intensive task and combining with standard clinical data adds additional layers of complexity
- Capturing the patient’s behavioral data through several sensors; their various social interactions and communications

The Institute for Health and Technology Transformation (Cottle, et al., 2013) distinguishes similar, but different challenges:

- **Industry readiness**
  The healthcare industry is unprepared to handle the deluge of data. They do not manage their data correctly. There is cultural reluctance to embrace big data
- **Data usability**
  Medical data is highly unstructured, making it hard to use. Trustworthiness of data is also an issue due to the human aspect of entering data
- **Data fragmentation and infrastructure**
  Data is stored in silo’s

- **Ownership, use, security**
  Who owns healthcare data, how and by whom can that information be used and for what purposes. Healthcare organizations must secure four types of data: personal information; clinical data; financial data and behavioral data.

Examples
There are numerous examples of big data analytics applied to healthcare. We considered 15 examples of big data analytics that are supposed to be examples of big data analytics in healthcare. However, upon closer examination only 8 of these 15 examples were actual big data analytics examples. More than half of the examples did not meet our definition of big data analytics. The results of this analysis can be seen in Table 2. More information on these examples can be found in Appendix A.

Name and reference	Big data?	Name and reference	Big data?
Heritage health prize (Heritage Health Prize, 2017)	No, low variety	Fraud prevention (Ottenheijm, 2015)	No, no variety
Project Artemis (Cottle, et al., 2013)	Yes	North York General Hospital (Raghupathi & V, 2014)	Yes
Joint-replacement (Cottle, et al., 2013)	No, no variety and no velocity	Care protocols (Raghupathi & V, 2014)	Yes
Google flu trends (Cottle, et al., 2013)	Yes	Brain injuries (Raghupathi & V, 2014)	Yes
Proper health (Ottenheijm, 2015)	Yes	Adverse drug effects (Raghupathi & V, 2014)	No, low velocity and low variety
Aurora health care (Ottenheijm, 2015)	Yes	Ebola control (Selanikio, 2016)	No
CPCT (CPCT, 2017)	No, low velocity	Parkinson's disease app (Selanikio, 2016)	No
Watson (Selanikio, 2016)	Yes		

TABLE 2 EXAMPLES OF BIG DATA ANALYTICS IN HEALTHCARE EVALUATED

3.3 DUTCH ACADEMIC HOSPITALS
This section discusses what is typical for the UMC's within the Dutch healthcare industry what the challenges are for big data analytics in this industry.

Background
Dutch academic hospitals have three main tasks: proving clinical care in the most complex cases, educate medical specialists (to be) and create new knowledge by doing groundbreaking medical research. The academic hospitals are responsible for one third of all research in the Netherlands (Chiong Meza, Van Steen, & de Jonge, 2014). They are financed by two ministries: Onderwijs, Cultuur en Wetenschap (Education, Culture and Science) and Volksgezondheid, Welzijn en Sport (Health) and furthermore receive funding for scientific research from third parties and health insurers pay the UMCs for care of patients.

As big data analytics in healthcare is supposed to change the industry, the academic hospitals in the Netherlands are the logical place to discover this potential. In an academic hospital there are researchers, medical specialist and the hybrid form: medical specialists that also do research. These different roles have different perspectives on care, which are visualized in Figure 5.

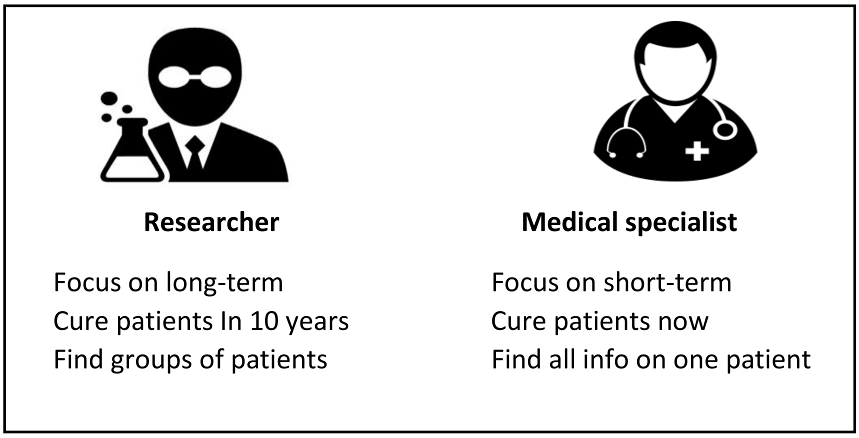


FIGURE 5 DIFFERENT VIEWS ON CARE AND USE OF DATA IN ACADEMIC HOSPITALS

Specific challenges for the Dutch healthcare industry

Nictiz (Ottenheijm, 2015) mentions challenges that are specific for the Dutch healthcare industry:

Technology

The traditional ICT architectures currently in use have a lot of limitations when it comes to big data. Some elements that are currently missing are the ability to store massive amounts of data, analyzing it and visualizing the outcome. Various data assets need to be combined for a thorough big data analysis means many parties have to share date with one another. Agreements need to be made on what technologies will be used.

Standardization

Currently there is almost no standardization for storing and exchanging healthcare data between healthcare providers. There are many different systems using different formats. Standardizing these formats and defining data definitions would be a step in the right way.

Data Access

A very important challenge to overcome is the fact that data is currently stored in a fragmented way. Every department within a hospital, every healthcare provider and even each patient could have its data stored in a silo, while big data initiatives can only succeed when data is combined. Although patients are more and more aware of the fact that they own their data, it is still a long way to go to change the mind-set towards a shared data-platform.

Privacy

Another very important discussion about applying big data to healthcare is ownership of information and who can use this information and for what purposes. Since the Electronic Patient File (EPD) has been discussed in the government, there has been a lot of resistance towards the project. Patients are wary of their data and breaches of their trust can be disastrous for big data initiatives. Some background information on the privacy concerns is provided below.

Privacy concerns in the Netherlands

In 2008 the Elektronisch Patiënten Dossier (Electronic Medical Filing system) was not accepted by the government because of privacy concerns throughout the country. There was a lot of negative publicity (Nu.nl, 2017) written about the EPD. Even though the EPD as it was proposed is no longer relevant, the privacy concerns are still present throughout the Netherlands.

Even today, Nictiz recognizes that privacy still is a great concern when discussing big data analytics in healthcare. The question that rises is who is owner of healthcare data and who can use this data and for what purposes. National and European laws are in place about the protection of healthcare data. Absolute control over privacy seems impossible: mostly because it is unclear who the owner is of the data and who can use this.

There is currently no national Electronic Medical Filing system in place. Every academic hospital does have an Electronic Medical Filing system, but it is not possible to easily share data with other healthcare providers in the Netherlands.

Ottes recognizes in his report for the ministry of Health (Ottes, 2016) similar challenges and concerns regarding big data in healthcare in the Netherlands. He mentions privacy and security of data. He also expresses concerns on the power of information: commercial companies are providing you with apps or websites in return for your data (e.g. Google). He wonders who has the authority for medical data. Besides these challenges, big data itself has limitations. It can show correlations but not if these are relevant, data can be manipulated and it is prone to biases. He considers it an extension to existing research methods and not a replacement of these methods.

3.4 EXPERT INTERVIEWS

Three experts on big data analytics in healthcare from outside the industry were interviewed. The purpose of these interviews was to find characteristics specific for the Dutch academic hospitals. They were very clear on privacy, data-driven use of big data and buy-in from the business. Their main messages are described below:

- Privacy is very difficult when it comes to big data. Privacy and security need to be considered because this is currently in its infancy. Privacy by design should be considered for all big data analytics attempts.
- Big data is best used when you do not have a use case to begin with. You should be able to have an open mind and use big data in a data-driven manner. A business case is not necessary to make big data analytics happen because of this data-driven approach.
- Technology is ready, and data-enthusiasts in hospitals might want to do big data analytics, but you first need to get support from the business. If the rest of the hospital does not want to go in that direction, progress towards a viable BD(A) usage will stifle.

3.5 CHARACTERISTICS

Based on the literature review on big data analytics in Dutch academic hospitals and the expert interviews we find the following characteristics that should specifically be addressed in the to be specified maturity model which will be discussed in the next Chapter.

These characteristics are:

- Standardization
- Data access

- Compliance with privacy laws and legislation
- Data-driven use of big data analytics
- Buy-in from the business
- Technology

Technology is added to the list of characteristics as it is a challenge specific for the healthcare industry. Even though it seems to be fundamental for big data analytics to exist, the existing ICT architectures are not sufficient for big data analytics in the academic hospitals. These characteristics will be used as the only requirements a maturity model on big data analytics in Dutch academic hospitals should meet, but as complementary requirements. The maturity model and these requirements will be discussed in the next chapter.

4. EXISTING MATURITY MODELS FOR BIG DATA ANALYTICS IN HEALTHCARE

In this chapter existing maturity models for big data analytics in healthcare are discussed. First, an introduction to maturity assessments is given. Then, the selection of maturity models to be assesses is discussed. Thirdly, the selected maturity models are compared on different aspects. Finally, a conclusion is drawn on the applicability of the selected maturity models to the Dutch academic hospitals.

4.1 MATURITY ASSESSMENTS

Maturity assessments are commonly used as an evaluative and comparative basis for improvement and in order to derive an informed approach for increasing the capability of a specific area within an organization (de Bruin, Rosemann, Freeze, & Kulkarni, 2005).

A big data analytics maturity model provides a systematic method for understanding existing big data analytics maturity (Brooks, El-Gayar, & Sarnikar, 2015). It includes a review of important business and technical processes, taking into consideration critical success factors for big data analytics within an organization.

Maturity models are valuable instruments for IT managers because they allow the assessment of the current situation of a company as well as the identification of reasonable improvement measures (Becker, Knackestedt, & Pöppelbuss, 2009).

A maturity model consists of a sequence of maturity levels for a class of objects. It represents an anticipated, desired or typical evolution path of these objects shaped as discrete stages. The bottom stage stands for an initial stage and the highest stage represents a conception of total maturity.

The highest level is what should be aimed for by users. However, the highest level should be *a perfect world* for big data analytics and will not always be reachable due to constraints such as costs.

One of the most referenced models that is still used as a base for many other models is the Capability Maturity Model (CMM) (Paulk, 1993). The CMM was superseded in 2002 by the Capability Maturity Model Integration (CMMI), which is a process improvement approach to aid organizations improving their performance. The CMMI consists of five maturity levels: initial, repeatable, defined, managed and optimized, which have been used in many other maturity models since then (de Bruin, Rosemann, Freeze, & Kulkarni, 2005).

Knight et al. (R, Knight, & Montgomery, 2012) state that the essential components of a maturity model are:

- **Levels**
  The transitional states in a maturity model
- **Model domains**
  Domains are a means for grouping like attributes into an area of importance for the subject matter and intent of the model
- **Attributes**
  Attributes represent the core content of the model grouped together by domain and level. They are typically based on observed practice, standards or other expert knowledge.
- **Appraisal and scoring methods**
  Appraisal and scoring methods facilitate the assessment using the model as a basis.
- **Improvement roadmaps**

In addition to being used for benchmarking, maturity models can be used to guide improvement efforts.

4.2 EVALUATING EXISTING MATURITY MODELS

In this section, the assessed maturity models will be discussed and compared on various criteria.

4.2.1 ASSESSED MATURITY MODELS

A search on Google Scholar for maturity models results in many hits, touching upon various subjects. A search for big data analytics maturity models in healthcare does not give any results, so a broader search was necessary. We considered the widely accepted CMMI, but then found a version of the CMM specific for data management that is more relevant. Furthermore, we considered models that focus purely on big data analytics, models on business intelligence and data warehousing and one model on business intelligence in healthcare. Table 3 gives an overview of the found models, the scope of the model and origin, the maturity domains that are assessed in the method and the maturity levels.

Reference	Scope	Maturity domains	Maturity levels
Hortonworks (Dhanuka, 2016)	Big data maturity model, commercial	Sponsorship, data & analytics practices, technology and infrastructure, organization & skills, process management	Aware, exploring, optimizing, transforming
Commuzi et al. (Commuzi & Patel, 2016)	Big data maturity, scientific	Strategic alignment, data, organization, governance, information technology	5 levels, nameless
TDWI (TDWI Big data maturity model and assessment tool, sd)	Big data analytics maturity, commercial	Data management, analytics, governance, organization, infrastructure	Nascent, pre-adoption, early adoption, corporate adoption, mature
InfoTech (InfoTech, 2013)	Big data maturity, grey literature	People, process, technology, data	Explorer, analyzer, integrator, innovator
CMM Data management maturity model (Data Management Maturity, 2017)	Data management, grey literature	Data management strategy, data quality, data governance, data operations, platform & architecture, supporting processes	Initial, repeatable, defined, managed, optimizing
Gartner (Burton, 2007)	BI maturity model, commercial	People, processes & metric, technology	Unaware tactical, focused, strategic, pervasive
IDC (Magee, 2016)	Big data analytics, commercial	Vision, data, technology, people, process	Ad hoc, opportunistic, repeatable, managed, optimized
Lessanibahri et al. (Lessanibahri, Gastaldi, Pietrosi, & Corso)	BI in healthcare, scientific	Functional, technological, diffusional, organizational	4 levels, nameless

TABLE 3 OVERVIEW OF THE ASSESSED MATURITY MODELS

Becker, Knackstedth and Pöppelbuss describe an evaluation method to review maturity models. In this method they distinguish several criteria to evaluate on: comparison with existing maturity models, iterative procedure, evaluation of method, multi-methodological procedure, identification of problem relevance, problem definition, targeted publication of results and scientific documentation. These criteria contribute to a methodically well-founded maturity model. The assessment of the models on the first 8 requirements is visible in Table 4. Requirement 8, scientific documentation, was specifically left out because Table 3 already shows the nature of the maturity models in the second column (scientific, commercial or grey literature). Besides these requirements, we also check domains and attributes of the maturity models and compare these to the characteristics of the industry described in Chapter 3.5.

Criteria used for evaluation defined by Becker et al.

1. Comparison with existing maturity models
   The need for the development of a new maturity model must be sustantiated by a comparison with existing models.
2. Iterative procedure
   Maturity models must be developed iteratively.
3. Evaluation
   All principles and premises for the development of a maturity model must be evaluated iteratively.
4. Multi-methodological procedure
   The development of maturity models employs a variety of research methods e.g. with interviews and a literature review.
5. Identification of problem relevance
   The relevance of the problem solution proposed by the projected maturity model for researchers and/or practitioners must be demonstrated.
6. Problem definition
   The prospective application domain of the maturity model, as well as the conditions for its application and the intended benefits, must be determined prior to design.
7. Targeted presentation of results
   The presentation of the maturity model must be targeted with regard to the conditions of its application and the needs of its users
8. Scientific documentation
   The design process of the maturity model should be documented in detail

Requirement	Comparison with existing models	Iterative procedure	Evaluation	Multi-methodological procedure	Identification of problem relevance	Problem definition	Targeted publication of results
Hortonworks	-	-	-	-	Strategic roadmap needed for big data challenges	Assess big data capabilities	Documentation and assessment tool
Commuzi	Comparison with 7 other models	Multiple iterations with case studies	Expert feedback with surveys	Literature review, expert interviews, case studies	Need for theory and tools to support ability to leverage big data	Help leverage big data and appropriate value derived from it	Scientific report, model
TDWI	-	-	-	-	Understand how big data deployments compare to peers	Determine maturity of your organizations big data initiatives	Free guide, assessment tool and webinar
InfoTech	-	-	-	-	-	Determine what data should be used in big data initiatives and what insights can be gained	Assessment tool
CMM	-	-	-	-	Need for model to evaluate data management capabilities	Reference model for process improvement for data management	Documentation, Assessment tool and guidelines for $100
Gartner	Comparison with older versions of Gartner model	-	-	-	Focus from IT is needed to keep BI connected to analytics	Take a strategic approach to BI, analytics and performance management	Documentation, assessment tool, guidelines, interpretation guidelines for subscribers
IDC	-	-	-	-	-	Help CIOs align their business value goals with IT strategy	Documentation of the model and an assessment tool
Lessanibahri	-	One iteration	Expert feedback with votes	Literature review, case study	Healthcare organizations fail in extending BI suites from pilots to larger domains	Measure and increase the maturity of BI solutions in healthcare organizations	Maturity model, scientific report

TABLE 4 COMPARISON OF MATURITY MODELS ON SCIENTIFIC DESIGN

Table 4 shows that only the maturity model from Commuzi et al. and Lessanibahri et al. meet most of the requirements set by Becker et al. Because Commuzi et al. consists of five maturity levels, which is in agreement with the most adopted maturity model, the CMMI, and it is the only model that actually meets all requirements, we will use this as a base to compare the rest of the models to in the next sections.

4.2.2 COMPARING DOMAINS

In the previous section it was decided that the model from Commuzi et al. will be a base to compare the rest of the models against. Because Commuzi et al. describe in their discussion that the domain privacy should be added in future research, this was also used in the mapping. All domains of the models were mapped against the domains used in the model from Commuzi et al. The results of this mapping are shown in Table 5. The domain privacy is not mentioned in any of the models. Some of the models do consider privacy as an attribute of a dimension. Strategy and governance are not always mentioned as domain. Technology, data and organization are always present although the precise terms may vary.

Model	Strategy	Governance	Technology	Data	Organization	Privacy
Hortonworks	Sponsorship	Process management	Technology and infrastructure	Data and analytics	Organization and skills	-
Commuzi et al	Strategic alignment	Governance	Information technology	Data	Organization	-
TDWI	-	Governance	Infrastructure	Analytics, Data management	Organization	-
InfoTech	-	Process	Technology	Data	People	-
CMMI DMM	Strategy	Governance	Platform & Architecture	Data quality, operations	Supporting processes	-
Gartner			Technology	-	People	-
IDC	Vision	-	Data technology	Process	People	-
Lessanibahri et al	Diffusional	-	Technological	Functional	Organizational	-

TABLE 5 COMPARING DOMAINS FROM DIFFERENT MATURITY MODELS

4.2.3 COMPARING ATTRIBUTES

All models have different attributes, but sometimes attributes with differing names are similar. To compare the coverage of all the models, a mapping was made of all attributes that are used in the models, grouped per domain. The results can be seen in Table 6. There is a lot of overlap in the general attributes, but the privacy domain is disregarded in many models. Another interesting attribute is the presence of the business case. The experts that were interviewed explicitly stated that a business case should not always be necessary (See Chapter 3.4).

Category	Hortonworks	Commuzi et al	TDWI	InfoTech	CMMI DMM	Gartner	IDC	Lessaibahri et al
Strategic alignment								
Big data strategy	x	x	x	x	x	x	x	x
Funding	x	x	x	x	x		x	x
Sponsorship	x	x	x	x		x	x	x
Business case	x		x	x	x	x	x	x
Performance management	x	x		x		x	x	x
Big data roadmap		x	x			x		
Data processing SLAs		x	x			x		
Governance								
Big data steering committee	x	x	x	x				
Big data policies	x	x	x			x	x	x
Central data definitions		x	x	x	x			
Data stewards		x	x		x	x		x
Monitoring			x		x			
IT								
Big data storage	x	x	x	x	x	x	x	
Cloud storage	x		x			x		

Analytic tools	x	x	x	x	x	x	x	
Centralized infrastructure	x	x	x		x	x	x	x
Infrastructure strategy			x					
new IT on legacy systems			x					
Data access				x			x	x
Data								
Data collecting	x	x		x		x		x
Data storage	x	x	x					
Data integration	x		x	x	x	x		x
Data analysis	x		x	x	x			x
Data types	x	x	x	x	x	x		
Data sources			x		x			
Data usability			x		x		x	x
Data quality			x	x	x			x
Privacy								
Privacy policies		x	x	x			x	
Data access			x	x			x	
Data encryption			x	x				
Anonymization			x					
Compliance with law				x				
Risk management					x			
Organization								
Big data Analytical skills	x	x	x	x		x	x	x
In-house skills	x			x		x		x
Big data analytics team	x					x	x	x
Usage of bi data analytics		x	x	x	x	x	x	x
Attitude towards big data		x	x	x	x			
Education of technical people			x	x		x	x	x
Training budget				x			x	x

TABLE 6 COMPARING ATTRIBUTES FROM DIFFERENT MATURITY MODELS

4.2.4 COMPARING MATURING METHODS AND SCORING

In this section a comparison is made of the way an assessed organization would mature and how a score is obtained.

The CMMI is very clear on how to mature: you need to meet all set practices or activities in order to mature. The model is not suited to visualize nuances. This makes it clear in terms of scoring. The model does not address big data specifics.

Hortonworks offers a very clear representation of the as-is maturity and the to-be maturity. Is very clear, with a questionnaire that contains 80 questions in total. Nuances in maturity can be made, although it does not offer insights on attribute level. The model is not made for healthcare specific.

Commuzi et al defined a maturity model that has clear descriptions of where an organization should be for every domain on each maturity level. However, it does not provide a clear questionnaire. The model is made to measure big data maturity, but is not specific for healthcare.

TDWI offers a model that can show nuanced in maturity per dimension. It can show more details per maturity dimension, but does not offer one overview of the maturity that immediately offers insight into the maturity. It does provide a clear scoring method that is easy to use and understand. This maturity model is not made to fit the healthcare industry.

InfoTech offers a big data maturity model with a questionnaire that immediately shows the current maturity. It does not offer future recommendations or a graphical representation thus is not as suited for communication. The model is not specific for the healthcare industry.

Gartner provides a very well documented maturity method that offers a lot of documentation. The model is very extensive but especially the graphical representation is lacking because it does not provide any details on the assessed maturity. It would be difficult to compare different assessments with each other without having to dive deep into the model. The model is made to measure business intelligence maturity in general.

IDC does not have a clear user manual or a graphical representation. Besides some global information on the model such as domains, attributes and maturity levels, it does not offer deep insight into the model. This makes it unusable for end-users to use by themselves. The model measures big data maturity, not specifically for the healthcare industry.

Lessanibahri et al. developed a maturity model specifically for business intelligence in healthcare. The questionnaire that they developed is not as easy to understand as others. However, it is the only model that shows interdependencies between different attributes. The graphical representation offers insight on attribute level, but does not provide an overall perspective on the maturity.

4.2.5 Checking characteristics of the industry

In Chapter 3.5 the characteristics of the industry (Dutch academic hospitals) were discussed. These were:

- Standardization
- Data access
- Compliance with privacy laws and legislation
- Data-driven use of big data analytics
- Buy-in from the business
- Technology

In Table 7 the maturity models are assessed on these five characteristics. No model meets all characteristics from the industry. Data-driven use of big data analytics is not explicitly described in any of the maturity models.

Characteristic	Standardization	Data access	Privacy	Data-driven use	Buy-in from business	Technology
Hortonworks					X	X
Commuzi et al	X		X		X	X
TDWI	X		X		X	X
InfoTech	X	X	X		X	X
CMM	X					X
Gartner					X	X
IDC		X	X		X	X
Lessanibahri et al		X			X	X

TABLE 7 COMPARING MATURITY MODELS ON INDUSTRY CHARACTERISTICS

4.3 CONCLUSION

In the previous sections comparisons of the maturity models were made on their scientific relevance, domains, attributes, maturing methods, scoring methods and the presence of characteristics specific for the industry as determined in chapter 3.5. There is not one of the models that scored well in all these different comparisons. The model from Lessanibahri et al. is the only model specifically made for healthcare, but the model from Commuzi et al. scored best on scientific relevance. The model from TDWI and InfoTech considered privacy attributes, but these are both commercial models where the design phases of the models is not described anywhere thus it is not possible to use these as a base for further research. None of the models has privacy as a domain, which is a characteristic of the domain according to the expert interviews. Only Commuzi et al. mention explicitly that the privacy domain should be looked into for future research. There is not one method with a scientific description of the model, the maturity model, an assessment method and an assessment guide. Most importantly, none of these models are specific for big data analytics in healthcare.

So the conclusion is that the assessed maturity models do not fit the scope of this research and subsequently a new maturity model specific for the Dutch healthcare industry was developed. However because the models from Lessanibahri and Commuzi came closest to meeting all requirements, they will be used as the starting point for the new model. The new model will be discussed in the next chapter.

5. BIG DATA ANALYTICS MATURITY MODEL FOR DUTCH ACADEMIC HOSPITAL

In this Chapter, the developed maturity model is discussed. The purpose, structure, content and instructions of the model are discussed. Finally, the adjustments made to the model after each iteration are discussed.

5.1 PURPOSE OF THE MODEL

This model should be useable by employees in an academic hospital with a role in big data analytics and external assessors. It needs to strike an appropriate balance between the reality and model simplicity. A model that is oversimplified may not reflect the reality and may not provide sufficient meaningful information for the audience. But a model that is too complicated may have the potential of incorrect use and may lead to wrong conclusions.

The purpose of the model is for Dutch academic hospitals to be able to assess their big data analytics maturity and be able to compare this to other academic hospitals. The model will be domain specific: it should capture the complexity of big data analytics in Dutch academic hospitals. It will adhere to the previously stated characteristics specific for this industry.

5.2 STRUCTURE OF THE MODEL

In the previous paragraph, it was mentioned that users of the model should be able to assess the big data maturity of an academic hospital in the Netherlands. Several models were discussed and compared. The model from Lessanibahri et al. and the model from Commuzi et al came closest to meeting all requirements. These models were used as a starting point for the design of the model. The positives of the other models were also used as an interpretation for the design.

5.2.1 MATURITY LEVELS

This model consists of five maturity levels, based on the CMMI. These five levels are initial, repeatable, defined, managed and optimized. The interpretation of the levels was made with the input of one of the experts that was interviewed (see Chapter 2.4). The main characteristics of these levels are described:

Initial
This is the starting point for big data analytics. Big data analytics is not performed within the academic hospital. There is no strategy on big data analytics and there are no systems or protocols in place to support this.

Repeatable
Big data analytics is not yet something that is embedded within the organization or standard, but there are small ad-hoc projects that can serve as a proof of concept within the hospital. Big data analytics could be repeated, but there is no standard established. It can still be a time consuming process to gather data.

Defined
Big data analytics can be performed within the hospital and there is a standard business process. There are processes in place to facilitate big data analytics throughout the hospital, but these still require manual labor and can thus be time consuming.

Managed
Big data analytics are really embedded in processes of the academic hospital. Next steps are taken to prepare processes, protocols and IT to enable big data analytics on a national level.

Optimized
Big data analytics have reached the highest level of maturity for academic hospitals in the Netherlands. It is possible to perform big data analytics on data generated by different healthcare providers. The data is findable and accessible, without having to go through manual processes that might be time-consuming. There is one process that is adopted by all academic hospitals and only one IT system supporting this. Importantly, big data analytics can be performed proactively and not only retrospectively. This means that analyses are not only focused on proving a suspected theory or relation, but rather to discover new insights in the data.

5.2.2 Representation of the model

Representation of the model is very important (de Bruin, Rosemann, Freeze, & Kulkarni, 2005). It is widely accepted, and a basis for assessment in many tools, to use one-dimensional linear stages. This results in an average maturity stage being provided for an entire entity. This does not adequately represent maturity within complex domains, providing little guidance to an organization to improve the current situation.

An alternative would be the 'stage-gate' approach. A stage-gate approach offers more differentiated maturity assessments within complex domains. This is achieved by providing an additional layer of detail that enables separate maturity assessments for a number of discrete areas. These layers can be represented by the domain and its attributes. This method will give an organization a deeper understanding of their relative strengths and weaknesses in the domain.

Because the goal of this research is to provide an actionable model that provides concrete advice on this matter, a stage-gate approach seems appropriate. With this method, an academic hospital would not be at maturity level 2 or 3, but each domain could be at another level of maturity. A matrix representation is used to make sure that a model will give insight in the current situation in one overview. Besides the overview of the maturity, a representation of the maturity of each attribute is also provided. Maturity models are not only used to assess the current maturity of an academic hospital, but can also be used as a benchmark for a UMCs progression. Therefore both representations also show the 'to-be' maturity in two years from now, following the roadmap already laid out. Examples of these representations are visible below.

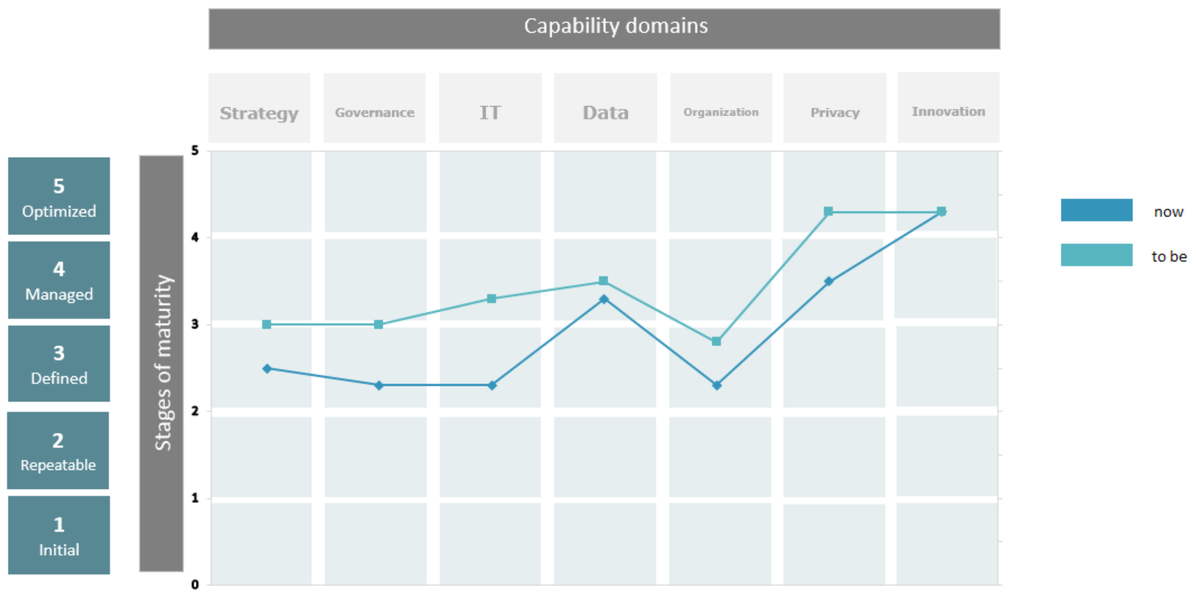


Figure 6 Big data maturity overview – example

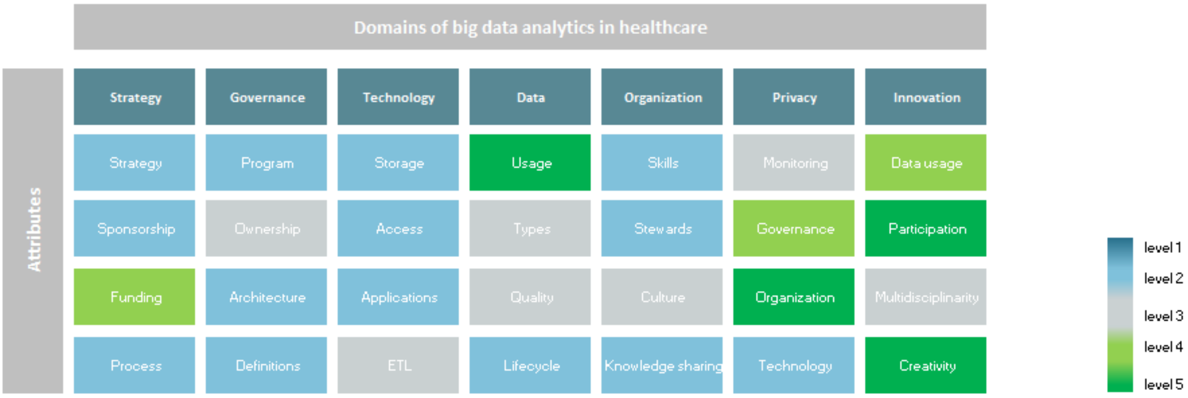


FIGURE 7 BIG DATA MATURITY ON DETAILED LEVEL NOW – EXAPMPLE

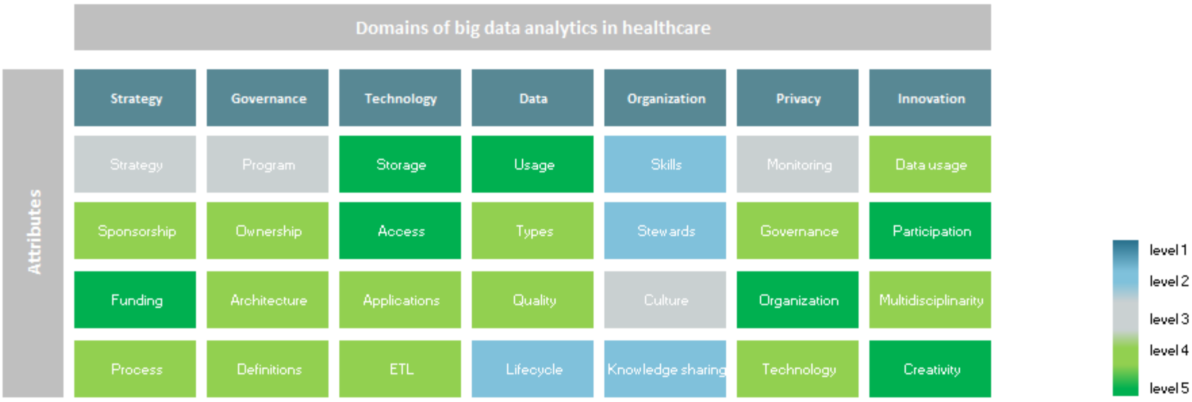


FIGURE 8 BIG DATA MATURITY ON DETAILED LEVEL TO BE - EXAMPLE

5.2.3 SCORING

Every domain consists of several attributes. Each of these attributes is represented with one question with five answers. The first answer corresponds to maturity level 1, the second answer corresponds to maturity level 2, etcetera. Answering a question with answer 1 results in 1 point, answering a question with answer 2 results in 2 points, and so on. When all questions are answered, the maturity of the hospital can be determined.

For every domain, add the points obtained by the questions of all attributes in that domain. Then, divide the sum of these points by the number of attributes in that domain. This figure, rounded to one decimal, is the current level of maturity for that specific domain. Repeat this for all domains. Finally, all maturity levels are added and divided by the number of domains. This figure, rounded to one decimal, is the overall current maturity of the academic hospital.

This is repeated for the to-be maturity score of the academic hospital.

5.3 CONTENT OF THE MODEL

In this subsection the content of the model is discussed. The process of creating the content is first discussed and then the content itself is discussed. First, the domains and the attributes of the model are described followed by the attributes. Finally, the questions of the model are discussed.

5.3.1 Determining domains and attributes

To create the model, the domains that were used by Commuzi et al. were adopted. The Privacy domain was added because it is an important characteristic for the health industry and Commuzi et al. mention in their conclusion that this should be added in a next phase. Furthermore, an Innovation domain is added because it is a characteristic to use big data for a data-driven approach instead of hypothesis-driven.

Table 6 was used to determine the attributes for each of the domains. The most used attributes were picked first and then discussed with the experts (See Chapter 2) to find the first set of domains and attributes for the model.

The definition of the Privacy domain and the attributes it consists of were derived from a maturity model to assess information security governance (Williams & Andersen, 2001). One expert on IT security with over 20 years of experience (see Appendix C) was consulted to discuss the domain and the attributes.

The definition of the Innovation domain and the attributes it consists of were derived from articles on team innovation in healthcare (Fay, Borrill, Amir, Haward, & West, 2006), creative thinking and data analysis (Sutherland), data-driven cultures (Harvard Business Review, 2012) and big data teams (Saltz & Shamshurin, 2016). An expert on innovative cultures in academic hospitals was consulted to discuss the content of the domain, see Appendix C.

After feedback from the experts on big data analytics and the case studies via surveys, as described in Chapter 2, the model was iteratively adjusted. The final version of the model is discussed below. Previous versions of the model can be found in Appendix B.

5.3.2 The domains of the model

The model consists of seven domains that altogether contribute to big data analytics in academic hospitals. These domains are Strategic Alignment, Data Governance, Information Technology, Data, Organization, Privacy and Innovation. Each of these domains consists of domain specific attributes that contribute to making a domain comprehensible and measurable.

1 Strategic alignment
Without the support of the full organization, big data analytics will not succeed. The board has to set out a strategy that will define *how* big data analytics will be used within the organization. This should be defined in a strategic document. A clear big data strategy is considered as key to successful adoption of big data analytics within an organization. This strategy has to be adopted by the whole organization to make this succeed. The strategy needs to formulate a clear vision, obtaining the buy-in within the whole organization and not only IT. Sponsorship also involves funding and an advocate of the program in the board of the hospital.

2 Governance
Big data governance formulates policies relating to optimization, privacy and monetization of big data by aligning the objectives of multiple functions. These policies are on metadata (setting definitions for data), access (who gets access to data?), data ownership, data quality, data security, data assets and data lifecycle. Governance does not consider data on operational level but sets guidelines and rules on how to use the data, and who is responsible for what in the organization considering big data analytics. Data ownership should be defined for every data source at each point in the big data analytics process.

3 Information technology
Big data analytics is in essence a technological solution. In the perfect world, all data sources should be connected to a central data warehouse. This includes all internal data sources, but also applications that are used by patients or medical monitoring devices such as heart rate monitors. To ensure the best care for the patient, all relevant patient data should be available. A central data storage facility should be able to deal with the volume, velocity and variety of big data.

4 Data
Data is a domain that is also, besides technology, part of the core of big data analytics within the hospital. To get the best result out of big data analytics, data should be of high quality. Data can be used for many different purposes. Five ways of using data are distinguished, from low to high maturity, reporting, monitoring, evaluation and finally prediction. All data should be available for data analysis. So not only structured data, but also unstructured data and (near) real-time data. Besides the internal data sources of the UMC, external data sources should also be considered such as data on air pollution.

5 Organization
A big data analytics project will only succeed when the right people are hired or trained to do this job. The organization should be as digital as possible. Big data analytics can only be performed on digital data. If most data is still in paper documents, this data cannot be used. The industry is still hesitant towards big data analytics as it is not the adopted way of doing research, which is hypothesis-driven. The adoption of data-driven decision making is considered in this domain. Finally, academic hospitals are almost a small town on their own. They all employ around seven to ten thousand employees. In an ideal world, these employees would share their knowledge on big data analytics so that the wheel is not reinvented again.

6 Privacy
Privacy issues have become increasingly urgent as more and more personal data is online. The electronic patient files are special because there are concerns of patients about disclosure of personal health information to third parties such as insurers or employers. However, the data should be shared and not kept within silos, so there has to be a mutual understanding between data sharing while keeping privacy and security standards high. Ideally, the data in a central data storage that is used for analytics will never leave a secure environment such as a so-called *sandbox.* Security audits should be in place and there should be frequent checks to see if the situation meets compliance and legislations standards. This also involves training users of this data on the matter. Data should be shared, but not without protection. Also, as patients should be notified in case of important findings, pseudonymization should be used.

7 Innovation
The last domain of the model is the innovation domain. As mentioned earlier, big data analytics can be useful to discover new relations or theories within the data that can be found by exploring the data in a different way than the data is now mostly used: starting with a hypothesis and proving it with the data. Innovation is closely related to a climate of creativity and the composition of teams. Multidisciplinary teams have a higher capability of thinking outside the box.

5.3.3 Attributes of the model
In this section, the attributes of the model are discussed per domain. For each of these attributes, the ideal situation is described, corresponding with the highest level of maturity.

STRATEGIC ALIGNMENT

The Strategic Alignment domain consists of four attributes: Strategy, Sponsorship, Funding and Adoption.

Strategy
Big data analytics should be considered for the academic hospital's strategy and documented in an actionable way i.e. with a roadmap. Besides a strategy for the UMC, there should be one actionable, documented strategy for big data analytics between UMCs.

Sponsorship
A sponsor promotes big data analytics and grants the mandate for the program. The sponsor continuously stretches the programs importance. Someone is considered a sponsor of big data analytics when this is (explicitly or implicitly) embedded in the description of their role. Besides the CEO of the academic hospital, there should also be a sponsor for big data analytics on national level.

Funding
The funding for big data analytics programs should come from business units and not from the IT department. There should be a shared budget between healthcare providers to support big data programs.

Adoption
The adoption-rate of big data analytics in the hospital is measured. Big data analytics should be used throughout the UMC, all using the same protocols and guidelines to structure the analysis. In an ideal situation, this should be true for all UMCs, all using the same protocols.

DATA GOVERNANCE

The Data Governance domain consists of four attributes: Program, Ownership, Architecture and Definitions.

Governance program
Somewhere in the organization, one or more groups are authorized to make rules and key decisions. This group of individuals (or a hierarchy of groups) typically represents a cross-section of stakeholder groups. Together, they define a set of rules in the form of policies, standards, requirements, guidelines, or data definitions. This group of rule-makers may go by several names: a Data Governance Board, a Data Stewardship Council, a Data Governance Program, etc.

There should be a formal data governance program in the UMC, but ideally one formal program for all UMCs in the Netherlands.

Data ownership
Even though applications are often owned by IT, data ownership is often undefined. The owner of a business process is often also the owner of a related data asset. An owner has the highest level of responsibility over a specific data asset. Data ownership should be clear for all data in the UMC, and for all data shared between UMCs.

Data architecture
A data architecture should provide a clear overview of all data that is currently available for big data analysis. In the ideal situation not only data sources are known, but also metadata on this data. The data architecture should at least be available for the central data storage such as a data warehouse or a data lake. A data lake is a shared data environment that comprises multiple repositories where users can access vast amounts of raw data. The differences between a data lake and a data

warehouse are that data lakes retain all data and not just data that is deemed relevant for reporting, it supports all types of data, it supports all types of users and it can easily adapt to changes. Ideally, all data in the UMC is findable, accessible, interoperable and readable (FAIR).

Data definitions
Data definitions should give information about data such as meaning, relationships to other data, origin, usage and format. There should be one standard set of data definitions used by all UMCs in the Netherlands for data in their central data storage.

INFORMATION TECHNOLOGY
The Information Technology domain consists of four attributes: Data storage, Data access, ELT and Applications.

Data storage
To make data access easier, and the data more manageable, data should not be stored in siloes. There should be one central data storage where all data sources are accessible and only one accepted infrastructure to access data from other UMCs.

Data access
Data access should not be difficult, or a lengthy process. Data access should not take more time than looking up what data to access, and this should not take time or intervention from the department managing the central data storage. Ideally, data access is an automatic process for all data sources in the hospital and all data that is shared between hospitals.

ELT
The ETL (extract – transform – load) process focus on extracting data from data sources, then transform the data to a certain set of rules or functions and then to load it into a central data storage which is often a data warehouse. However, with the rise of big data, these methods often take too much time and effort and are high maintenance. The solution for this, comes with the rise of the data lake and ELT. Extract all data from the data sources, store it raw into a data lake and then transform the data dependent of the request. When big data analytics reaches its highest level of maturity, a data lake is implemented and thus ELT is necessary. ELT should then be done automatically for all data in the data lake of the UMC and there should be one ELT standard to access other UMCs accessible data.

Some differences between ELT and ETL are shown in the table below.

Criteria	ETL	ELT
Time to load	Uses staging area and system, extra time to load data	All in one system, load at once
Time for transformation	Need to wait, especially for big data sized. As the data grows, the transformation time increases	All in one system, speed is not dependent on data size
Time for maintenance	High maintenance	Low maintenance
Implementation complexity	At early stage, requires less space and the results is clean	Requires in-depth knowledge of tools and expert design of the main large repository
Analysis & Processing style	Based on multiple scripts to create the views – deleting view means deleting data	Creating adhoc views – low cost for building and maintaining
Usability	Fixed tables, fixed timeline, used mainly by IT	Ad hoc, agile, flexible, usable for all purposes

TABLE 8 DIFFERENCES BETWEEN ELT AND ELT (ETL VS ELT: THE DIFFERENCE IS IN THE HOW, 2016)

Applications
Applications are necessary to perform big data analytics on. Preferably, there is one secured digital environment where data is analyzed. This ensures that data remains private and secure. There should be one digital environment with applications where data from multiple UMCs is analyzed.

DATA
The Data domain consists of four attributes: Data usage, Data types, Data quality and Data lifecycle.

Data usage
Data can be used for many different purposes: reporting, monitoring, evaluation and prediction. Ideally, big data analytics are used for prediction.

Data types
A big data analysis should not be limited because of the data types that are available. Not only structured data should be available, or aggregated data. Unstructured data, (near) real-time data such as monitoring devices and external data sources such as weather data should be available.

Data quality
If all analysts have to clean up their data set, this process will repeated many times. If the quality is improved at the original data source, following one standard for data quality used by all UMCs, this process will only have to be done once.

Data lifecycle
Just like any product, data typically goes through a number of stages. It is created, used, needs maintenance, back-ups, and eventually deletion or archiving. Lifecycle aspects should be considered on frequent basis for all big data sets used by analysts, within the UMC but also for the data sets with shared data from other UMCs.

ORGANIZATION
The Organization domain consists of four attributes: Skills, Digital, Culture and Knowledge sharing.

Skills
Big data analytics can only be done in the UMC if big data capabilities are present in the

organization. These skills should ideally be found in every department in the role description of an employee. But narrowing these skills to one department can be limiting, so these employees should be able to work on projects that have a wider scope than just that departments data.

Digital
A precondition for a high big data analytics maturity is how digital the academic hospital is. The UMC should be completely paperless and so digital that there is no longer any shadow-IT present. For example: the Electronic Patient Record systems are sometimes complex to work with. As a result, some information on patients is kept in Excel documents.

Culture
The adoption of big data analytics in processes in the hospital is an aspect of the big data maturity of the UMC.

Knowledge sharing
Within the academic hospital it is important to share knowledge on big data analytics. The UMCs are employing thousands of employees, so it can be difficult to find people working on big data analytics. Knowledge sharing on big data analytics should be done frequently within the UMC, and between UMCs.

PRIVACY
The Privacy domain consists of four attributes: Monitoring, Governance, Awareness and Pseudonymization.

Monitoring
Monitoring of compliance with privacy policies should be done for all data in the central data storage of the UMC. Privacy policies should be monitored automatically and flagged when necessary. There should be one privacy board responsible for monitoring privacy policies on data between UMCs.

Security
Being able to experiment with big data and queries in a safe and secure "sandbox" test environment is important to both IT and end business users as companies get going with big data. If data is only accessible via a secured “sandbox”, data privacy and protection can be provided.

Awareness
Not only monitoring and policies should be in place, the employees should also be aware of privacy and security issues of data. They should be trained for education purposes on privacy awareness of patient data. It should be on top of mind for all employees when dealing with data.

Pseudonymization
Data in the central data storage can only easily be used for big data analytics purposes, if it anonymized or pseudonymized automatically. Privacy by design should be standard for the UMC. This is an approach where privacy is taken into account in the complete engineering process. Ideally data from all UMCs can be shared while anonymization or pseudonymization of this data is done automatically.

INNOVATION
The Innovation domain consists of four attributes: Usage, Participation, Multidisciplinary teams and Creativity.

Usage
Big data analytics can be used to prove hypothesis or to use the data as a starting point and discover new ideas and theories. These methods can co-exist, but big data analytics should not only be used retrospective but also proactive. By combining the data with other data sources from third parties, the full potential of big data analytics is used.

Participation
Innovation should be part of the entire UMC. The focus should not only be internally, but also on innovation on a national scale. All UMCs are collaborating. However, not all are having a leading role in innovation programs on a national level.

Multidisciplinary
Big data analytics that are performed by multidisciplinary teams have a greater chance of finding innovative theories and unexpected results.

Creativity
The innovation team should have time to think creatively during their day-to-day conduct of work. This should be encouraged actively by their manager and the innovation team should be trained or educated on creative thinking techniques such as 'design thinking'.

5.3.4 QUESTIONS OF THE MODEL
In this section, the questions of the model are discussed. All the 28 attributes have one question with five answers to assess the maturity of the UMC on that particular attribute. An example of such a question can be seen below in Figure 9.

Question 1

Strategy

1 Big data analytics are not considered for the strategy of the academic hospital

2 There is awareness on the possibilities of big data analytics but this is not documented in a strategic document

3 There is a hospital wide documented big data strategy

4 There is a hospital wide documented big data strategy that is actionable, i.e. with a roadmap

5 Besides a hospital wide strategy on big data, there is a documented shared strategy on big data analytics between UMCs that is actionable, i.e. with a roadmap

Answer	2	**To be**	3

FIGURE 9 EXAMPLE QUESTION OF THE MATURITY MODEL

Each of the 28 attributes is defined by a question with 5 answers, each corresponding with one maturity level. Each question has five answers that progress stepwise in terms of maturity. The first answer corresponds to the first level of maturity and the fifth answer with the highest level of maturity. When answering the questions start with the first answer. If this answer fits the current situation best, stop there. Otherwise, progress to the next answer. Continue this progress until the best fitting answer is found. When no answer matches the current situation, choose the best fitting answer. The same should be done for the to-be situation in two years from now. Only specified plans should be taken into account, ideas on the future that are not yet agreed on should not be considered.

The questions are based on the attributes described in the previous section. For each of the attributes, the maturity level was used as a base to formulate the answer to the question. For example level 1 initial means that there is nothing happening with big data analytics in the entire UMC. This corresponds with the answer on the question for Strategy: "*Big data analytics is not considered for the strategy of the academic hospital*", see Figure 9.

5.4 INSTRUCTIONS

Instructions on how to use the maturity model will be discussed in this section. UMCs that want to assess their big data analytics maturity are advised to go through all steps described in this section.

Step 1: Role assignment
The hospital that wants to prepare for the assessment should first assign certain roles before it can start. One person should execute the assessment and define the employees that need to be involved, and one person high in the organization should sponsor the assessment.

Assessor
The assessor should execute the maturity assessment. The assessor needs to be familiar with big data and have enough experience within the hospital to know which persons to ask for the assessment. The assessor needs to collect data from IT and business. The assessor could be someone working on innovation or involved with research ICT. The assessor should have knowledge on all the different domains. The assessor does not have to be an employee of the academic hospital. This could also someone that is not working at the UMC but has experience in doing these assessments. Someone from outside the hospital could possibly do an assessment without being biased.

Sponsor
Since multiple people within the hospital will be asked to spend time on the assessment and the assessment might lead to change, someone high in the organization should sponsor the project. He or she should let the organization know that the project is important and that people should cooperate.

Step 2: Collecting the data
The next step is to start collecting the relevant data to answer the questions of the model. Roles that should be asked for input are (if present within the hospital):

Roles
IT architect
Head of Business Intelligence
Innovation manager
CIO
Privacy/ security officer
Researcher using big data analytics, or more advanced analytics in the hospital

The assessor should have enough knowledge of the hospital to know if these roles are present within the organization, and if not find suitable replacements.
Data can be collected by handing out the questions to the specific people chosen, or if there is not enough knowledge of big data analytics to let the questions directly be answered, interviews can also be sufficient. Evidence should also be collected. For example, strategies could be in place but the assessor should look for *evidence* that this strategy is actually executed.

Step 3: Determining the maturity
When enough data collected to answer all questions, the questions should be answered using the

scoring form, which can be found in Appendix E. When in doubt, the assessor should decide if more information is needed, or if he/she can make a decision himself. After finishing answering all the questions, the visualizations of the current maturity should be made. The questions should only be answered by the assessor.

A guideline to help answering the questions can be found in Appendix B.

Step 4: Interpretation
The model is a tool to quickly assess and illustrate the current maturity in big data analytics within the hospital. This is purely descriptive, and does not provide immediate improvement potentials. Therefore, interpretation of the outcomes is needed. There are a couple of steps to interpret the results:

1. As-is landscape
   The assessor should start by analyzing the current landscape within the organization. Define what the current limitations are, and how big data analytics is currently placed within the organization.
2. To-be landscape
   A hospital that wants to assess their current big data analytics maturity, most likely has goals concerning this topic. Make the to-be maturity clear for the academic hospital. For the to-be maturity a period of two years is advised.
3. Gap analysis
   Once both the as-is and the to-be landscape are clear, the assessor should analyze the gap between these landscapes. These gaps should be connected to domains of the maturity model.
4. Recommendations
   The gaps, and the related domains, should be translated by the assessor to actionable recommendations. This does not mean that the hospital should immediately strive for the highest maturity level, but to a reachable higher maturity that is aligned with the strategy of the hospital.

Step 5: Validation
To validate that the assessor collected the right information and interpreted the information in the right way, the assessor should make a report and/or presentation to all persons that were involved. An example of a report can be found in Appendix F. This should lead to a feedback session where the results might be slightly adjusted if needed. If a feedback session is difficult to organize, a survey could be sufficient. An example of the survey can be found in Appendix D.

After the validation, the report and/or presentation should be finalized. This report should explitily mention the evidence found that whas used for the assessment.

1. Table of contents
2. Introduction
3. Explanation of the model
    a. Maturity levels
    b. Domains and attributes
4. Research method
5. Results
    a. Strategic alignment
    b. Governance
    c. …
6. Conclusions and recommendations

FIGURE 10 EXAMPLE TABLE OF CONTENTS FOR THE REPORT

Step 6: Distribution of report

The final step of the maturity assessment is distribution of a report with the findings of the maturity assessment. Figure 10 shows a suggested table of contents for the report. After finishing the report, the assessor should actively spread the results within the UMC.

5.5 ADJUSTMENTS TO THE MODEL

This section describes the changes made to the model at the end of each iteration.

CHANGES MADE AT THE END OF THE FIRST ITERATION

The first version of the model can be found in Appendix B: The Maturity model. This version was evaluated by three experts on big data analytics, one person from Radboud UMC and three persons from VUMC. The survey was sent to all interviewees from the two case studies, but some interviewees from Radboud UMC felt they did not have enough expertise on big data analytics or maturity models to validate the construct of the model.

Survey responses

The results of the survey can be found in detail in Appendix D: Survey setup and products.

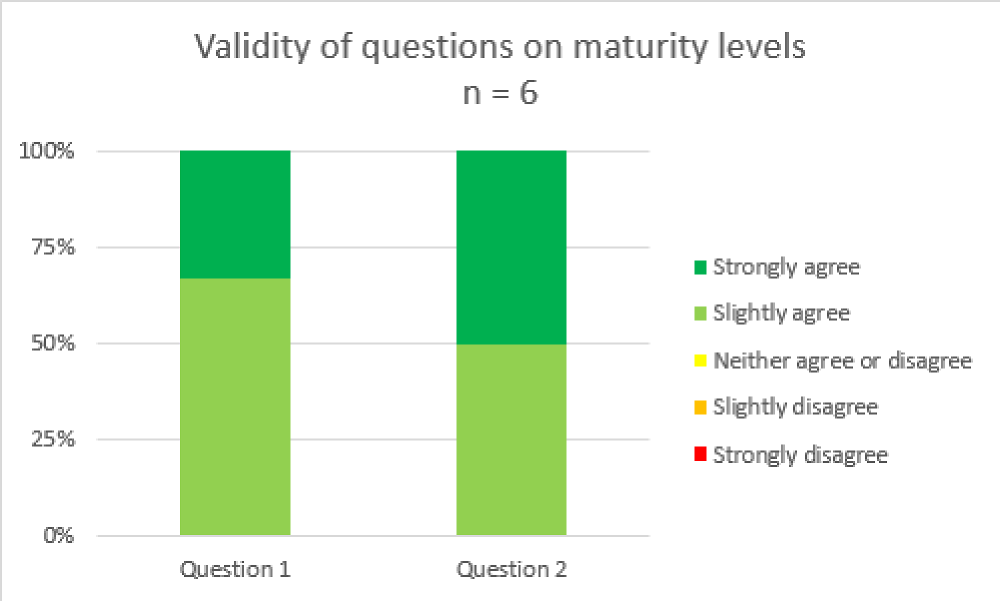


FIGURE 11 SURVEY ON VALIDITY OF THE MODEL VERSION 1.0 - ANSWERS TO QUESTION 1 AND 2

The questions on the maturity levels on sufficiency and accuracy where reflected positively. Most respondents slightly agreed that the maturity levels were sufficient to represent all maturity stages of the domain and they slightly or strongly agreed that there is no overlap between the descriptions of the maturity levels (Figure 11). No respondent would add any maturity levels and one respondent would add a definition of big data analytics in the documentation. Most respondents would not update the maturity level descriptions, except one respondent that described some improvement potentials as “sometimes the word protocols and sometimes the word processes is used and sometimes *discovering new insights* is used and sometimes *generate knowledge*. Use formulation of words more carefully”. One respondent suggested that the difference per maturity level for every attribute is not always in line with the descriptions of the model and should be reconsidered.

The questions on the domains and attributes were not received as well. The coverage, the mutual exclusion and the accuracy of the domains and attributes were questioned. The results from the questions on domains and attributes can be seen in Figure 12. The relevance of the attributes in a certain domain is agreed on with the author as 66% strongly agrees and 33% slightly agrees with the author. The coverage is questioned by one of the respondents, the rest slightly or strongly agrees that the domains and attributes cover all aspects involved in big data analytics in Dutch academic hospitals. Question 5 on mutual exclusion of domains and attributes has varying responses. 50% of the respondents strongly agrees with the author, while 16% slightly agrees, 16% neither agrees nor disagrees and 16% slightly disagrees. The accuracy of domains and attributes in their maturity level is questioned by one respondent that strongly disagrees. The other respondents either slightly or strongly agree.

No domains or attributes would be added or removed by the respondents. There were multiple suggestions to update or redefine the domains and attributes. The privacy domain has many attributes that are named the same as domains. This is considered confusing. The data stewardship attribute is suggested to not be a separate attribute, as data governance captures this attribute. The ETL attribute is suggested to be changed to ELT, as this is a concept that is mostly used with data lakes and big data instead of the more traditional ETL. The use of external data, not generated by the hospital, is suggested to address explicitly as this is a big opportunity of big data analytics. The privacy domain is suggested to be extended with attributes on security. One attribute is called culture which is a very big topic on its own. It is suggested to be very clear on this attribute as there can be many interpretations. Another suggestion is to put less emphasis on IT. One respondent felt that it should be measured how digital an academic hospital is, as data is necessary for big data analytics. Finally, a suggestion is made to involve the government's opinion on the topic.

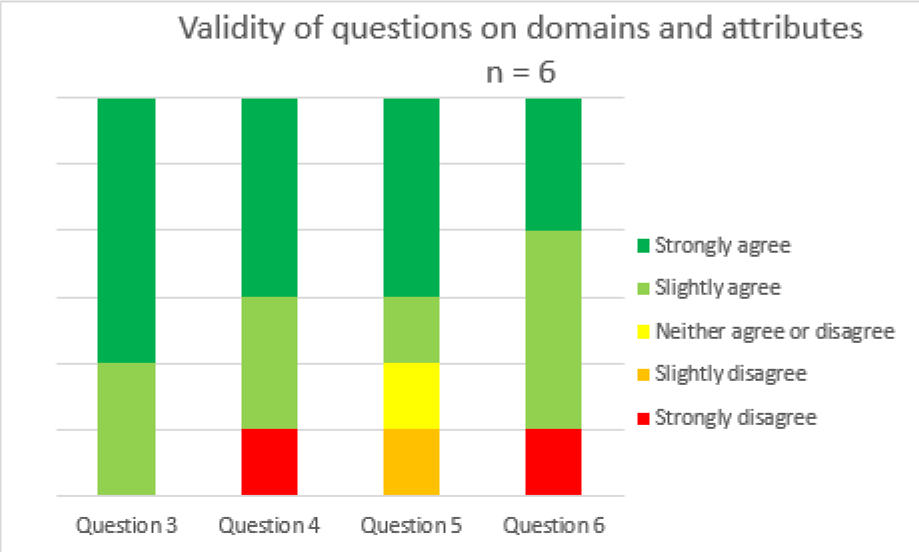


FIGURE 12 SURVEY ON VALIDITY OF THE MODEL VERSION 1.0 - ANSWERS TO QUESTIONS 3,4,5,6

Questions 7 till 10 are on the understandability of the maturity model. Results on those questions are visible in Figure 13. The understandability of the maturity levels is strongly agrees on by 66% of the respondents. 16% slightly agrees and 16% neither agrees nor disagrees. The assessment guidelines are received positively, 50% strongly agrees and 50% slightly agrees that these guidelines are understandable. The documentation has mostly positive feedback. 33% strongly agrees that the documentation is understandable and 50% slightly agrees. 16% neither agrees nor disagrees on the understandability of the documentation. As already visible in Figure 12, the domains and attributes are not all agreed on, question 10 also shows mixed responds. The understandability of why domains and attributes are assigned to a certain maturity level is strongly agreed on by 16%, slightly agrees on by 33%, neither agreed nor disagreed on by 33% and strongly disagreed on by 16%.

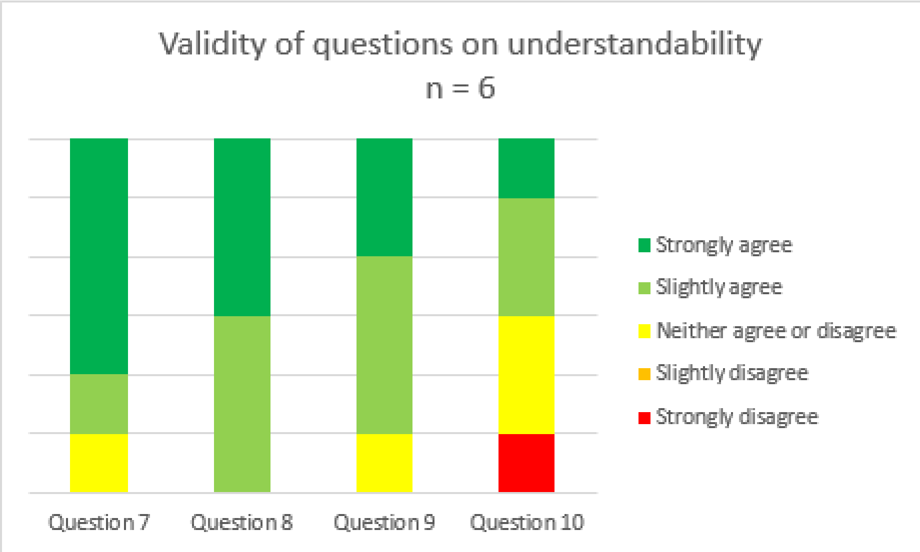


FIGURE 13 SURVEY ON THE VALIDITY OF THE MODEL VERSION 1.0 - ANSWERS TO QUESTIONS 7 – 10

Questions 11, 12 and 13 were on the ease of use of the scoring scheme, the assessment guidelines and the documentation (Figure 14). The scoring scheme is received as easy to use by 84%. 16% strongly disagrees on the ease of use of the scoring scheme. The ease of use of the assessment guidelines were strongly agreed on by 50% and slightly agreed on by 50%. The documentation was accepted by 84%. 16% neither agreed nor disagreed on the ease of use of the documentation of the model.

Some respondents suggested the scoring scheme could be improved by clicking on an answer instead of choosing from a drop-down menu.

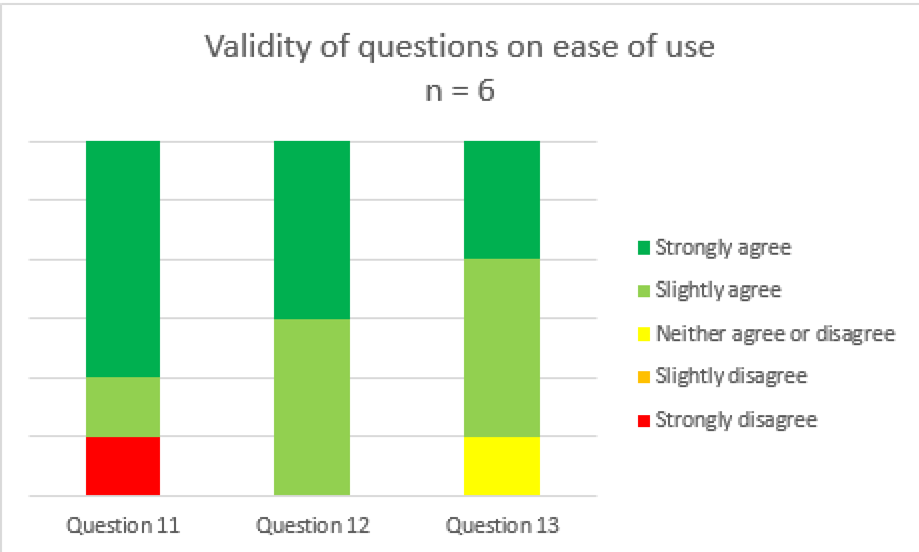


FIGURE 14 SURVEY ON THE VALIDITY OF THE MODEL VERSION 1.0 - ANSWERS TO QUESTIONS 11 - 13

The last two multiple choice questions were on the usefulness and practicality of the maturity model. The results are shown in Figure 15. 84% agreed that the maturity model is useful conducting assessments. 16% neither agreed nor disagreed. 84% agreed that the maturity model is practical for use in the industry, while 16% neither agreed nor disagreed.

Most respondents did not have tips to make the model more useful. One respondent mentioned that the model asks for a high level of expertise to use and thus not everyone will be able to use this. However, this is considered unavoidable because of the topic. The scoring scheme could be more useful by not using the different tabs and dropdowns. Another respondent recommended to make the report shorter by taking out the explanation of the model as this is already covered in the documentation.

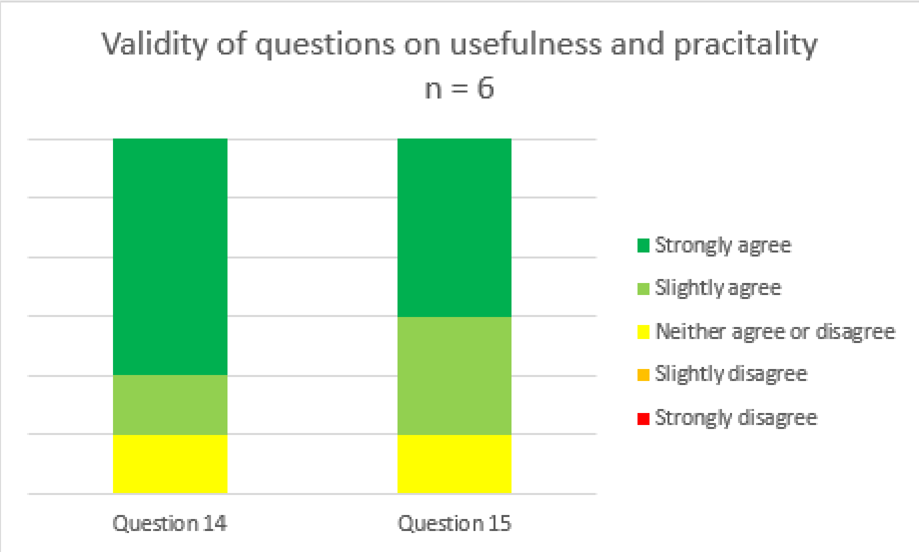


FIGURE 15 SURVEY ON THE VALIDITY OF THE MODEL VERSION 1.0 - ANSWERS TO QUESTIONS 14 AND 15

General comments on the model were that the CMMI is a good model to use as a basis for this maturity model. The domains are considered relevant for big data analytics in Dutch academic hospitals. Another respondent mentioned that some questions were not big data specific but more on IT maturity of the hospital. For example, an academic hospital can have a very high maturity in governance and privacy but not do anything with big data at all.

CHANGES MADE TO MODEL VERSION 1.0

The survey responses were used to make changes to the first version of the maturity model. The changes that were made were:

- A definition of big data analytics was added to the documentation of the model
- Many changes to answers of the questions of the maturity model to make them more clear and better fitting in the maturity level
- Some names of attributes were changed for clarification purposes. In strategic alignment, processes was adjusted to adoption. In governance, data governance program was adjusted to big data governance program. In privacy, Organization changed to Awareness and Technology changed to Pseudonymization
- The word data warehouse was changed to central data storage as a data warehouse is not general enough
- In the Information Technology domain, ETL was changed to ELT
- Organization has a new attribute, digital instead of data stewardship that considers how paperless an academic hospital is
- The interpretation of the attribute Culture in the Organization domain has changed to make the attribute more clear
- The Privacy domain has a new attribute Security instead of Governance.

The scoring scheme was specifically not altered because only two out of six responded that is was difficult to use. The other four respondents found the method easily usable.

The result of these changes is maturity model version 2.0. Both version 2.0 and version 1.0 of the big data maturity model are shown in Appendix B: The Maturity model.

CHANGES MADE AT THE END OF THE SECOND ITERATION

The second version of the model can be found in Appendix B: The Maturity model. This version was evaluated by two interviewees from the LUMC. The survey was sent to all interviewees from the two case studies, but some interviewees from LUMC felt they did not have enough expertise on big data analytics or maturity models to validate the construct of the model.

Survey responses

The results of the survey can be found in detail in Appendix D: Survey setup and products.

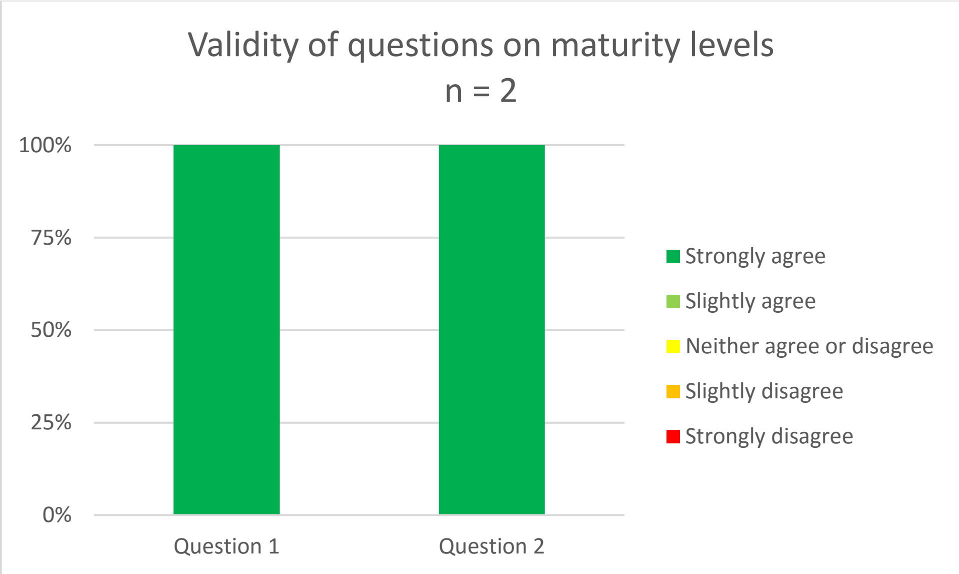


FIGURE 16 SURVEY ON VALIDITY OF THE MODEL VERSION 2.0 - ANSWERS TO QUESTION 1 AND 2

The questions on the maturity levels on sufficiency and accuracy where reflected positively. Both respondents strongly agreed that the maturity levels were sufficient to represent all maturity stages of the domain and they slightly or strongly agreed that there is no overlap between the descriptions of the maturity levels (Figure 16). No respondent would add any maturity levels. There was a question from a respondent on if machine learning should have a place in the model. The respondents would not update the maturity level descriptions.

The questions on the domains and attributes were not received as well. The results from the questions on domains and attributes can be seen in Figure 17. The relevance of the domains and attributes is strongly agreed on, as well as the coverage of the domains and attributes. However, one respondent slightly disagrees on the mutual exclusion of the domains and attributes. The respondents agree on the accuracy of the domains and attributes.

No domains or attributes would be added by any of the respondents. The domain Innovation is described as 'difficult. I do not understand the added value of this domain' by one respondent. The domains and attributes would not be redefined or updated by the respondents, but one respondent mentions that 'he could not always find a fitting answer'.

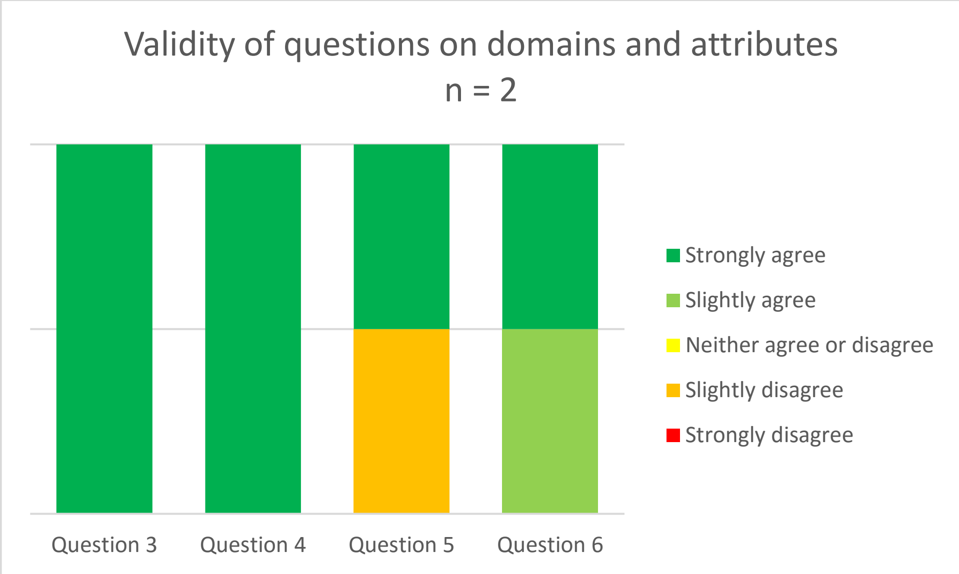


FIGURE 17 SURVEY ON VALIDITY OF THE MODEL 2.0 - ANSWERS TO QUESTIONS 3,4,5,6

Questions 7 till 10 are on the understandability of the maturity model. Results on those questions are visible in Figure 18. The respondents strongly agreed on the understandability of the maturity levels, the assessment guidelines, the documentation and that the domains and attributes are correctly assigned in their maturity level.

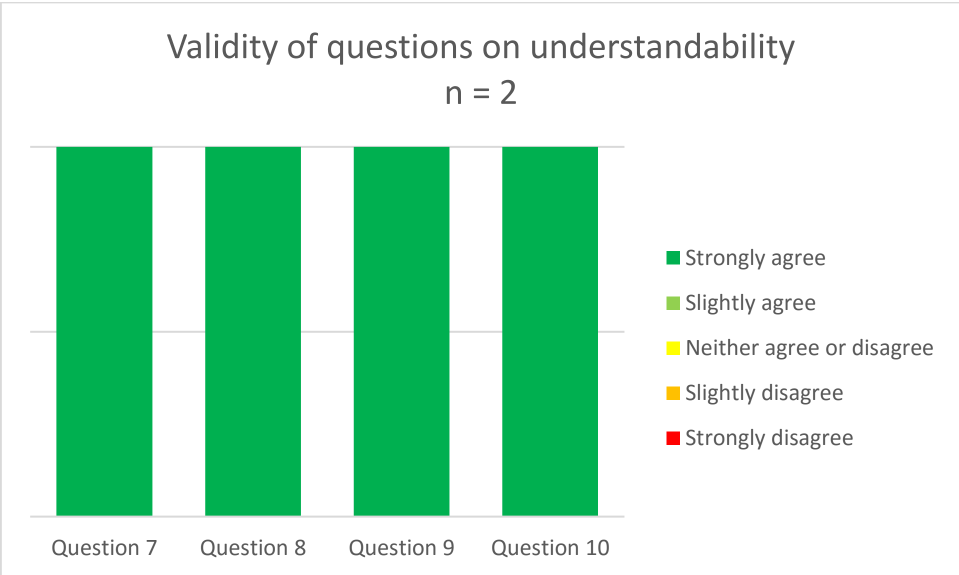


FIGURE 18 SURVEY ON THE VALIDITY OF THE MODEL VERSION 2.0- ANSWERS TO QUESTIONS 7 – 10

Questions 11, 12 and 13 were on the ease of use of the scoring scheme, the assessment guidelines and the documentation (Figure 19). Respondents strongly agreed on the ease of use of all these artefacts.

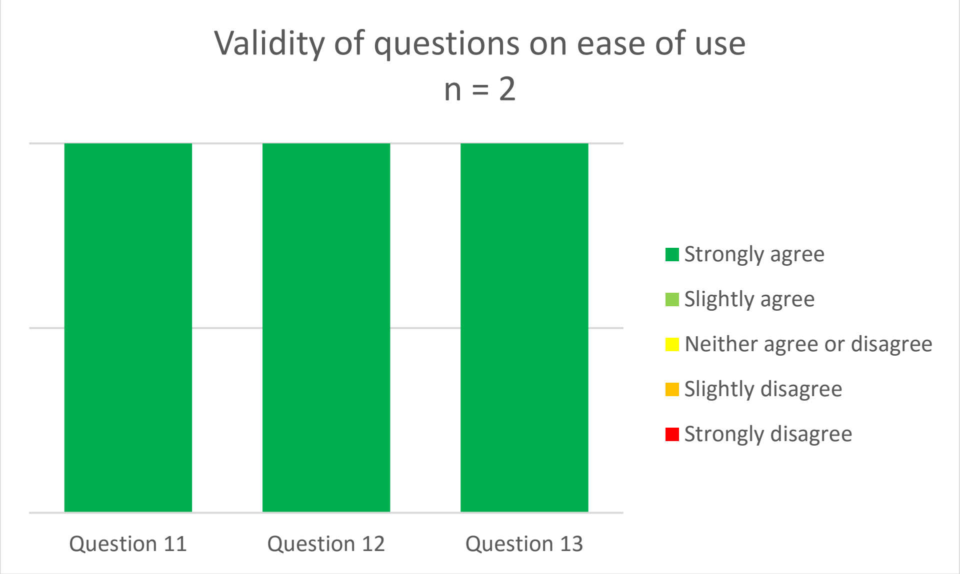


FIGURE 19 SURVEY ON THE VALIDITY OF THE MODEL VERSION 2.0 - ANSWERS TO QUESTIONS 11 - 13

The last two multiple choice questions were on the usefulness and practicality of the maturity model. The results are shown in Figure 20. The respondents strongly agreed on the usefulness of the maturity model for conducting assessments. One respondent strongly agreed that the maturity model is practical for use in the industry, and the other respondent answered that 'he did not know'.

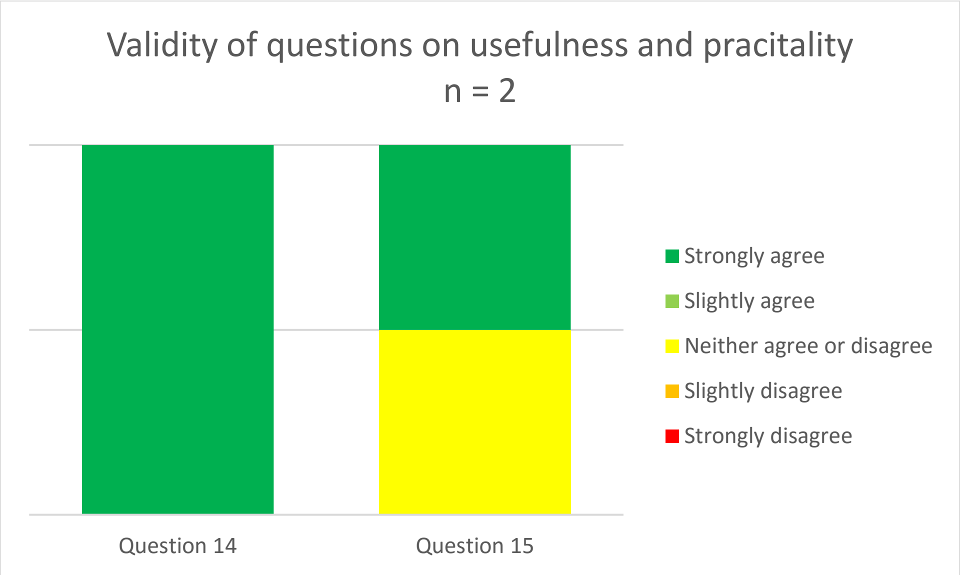


FIGURE 20 SURVEY ON THE VALIDITY OF THE MODEL VERSION 2.0 - ANSWERS TO QUESTIONS 14 AND 15

There was a general suggestion that setting the to-be maturity at only two years from now is restricting as the healthcare industry does not make many changes in such a short period of time. A suggestion is to change this into five years from now.

CHANGES MADE TO MODEL VERSION 2.0

The survey responses were used to make changes to the first version of the maturity model. The changes that were made were:

- Add to the assessment guidelines that the model should only be filled in by one assessor that has knowledge of all domains. Most respondents only have knowledge of one or several domains and should not be assumed to have sufficient knowledge to answer all questions.
- The attribute Anonymization was changed to Pseudonymization as this is more fitting for the industry

The to-be maturity is explicitly not changed from two years to five years because it is expected to be extremely difficult to make a prognosis for five years from now.

The result of these changes is maturity model version 2.1. Versions 2.1, 2.0 and version 1.0 of the big data maturity model are shown in Appendix B: The Maturity model.

6. RESULTS

In the first section of this chapter we will discuss the developed maturity model: acceptance, validation and gathering information. In the second section we will discuss the results from the case study at each studied academic hospital. In the third section we will discuss all the national initiatives that are working towards data sharing or big data analytics supported by the academic hospitals. Finally, we discuss the results of the sounding group sessions on the future of big data analytics in healthcare.

6.1 MATURITY MODEL

The maturity model was tested at three different case studies. The results from these case studies are discussed in the next section. The maturity model was developed to assess the maturity of big data analytics within academic hospitals in the Netherlands that should adhere to the previously stated characteristics specific for this industry.

Most aspects of version 1.0 of the maturity model were considered valid by the experts and the interviewees in terms of sufficiency, accuracy, relevance, mutual exclusion of content, understandability, ease of use and usefulness and practicality, as discussed in Chapter 5. Importantly, they found the model useful and practical to use in the industry. Version 2.0 of the maturity model was also considered valid by the interviewees.

However, it was clear that not all interviewees had sufficient expertise to answer all questions of the maturity method. They are mostly an expert in one or two of the domains, and sometimes lack knowledge to answer questions from other domains.

The maturity assessment is very time intensive. To get an overview of the current situation in an academic hospital is a difficult task as these organizations are huge, with nearly 10.000 employees each. Surveys alone are not sufficient to get an idea of the organization, interviews are necessary. Furthermore, the workload in these academic hospitals is high. Most interviewees do not have spare time for interviews or things other than their day-to-day job. This holds especially for researchers and doctors. As a result, the duration of one assessment can take up to two or three months.

6.2 CASE STUDIES AT UMCS

We discuss all the findings of the case studies at the UMCs in this section.

CASE STUDY AT RADBOUD UMC

For the case study at Radboud UMC, the maturity model version 1.0 was used. The as-is maturity and the to-be maturity of the Radboud UMC can be seen in Figure 21. The overall maturity of the Radboud UMC is 2,9 and the to-be maturity of the Radboud UMC is 3,5.

Some domains are much more mature than others. The domain Innovation and Privacy are much more mature than the five other domains. The domain data is slightly more developed, on level 3.The other four domains are all on level 2.

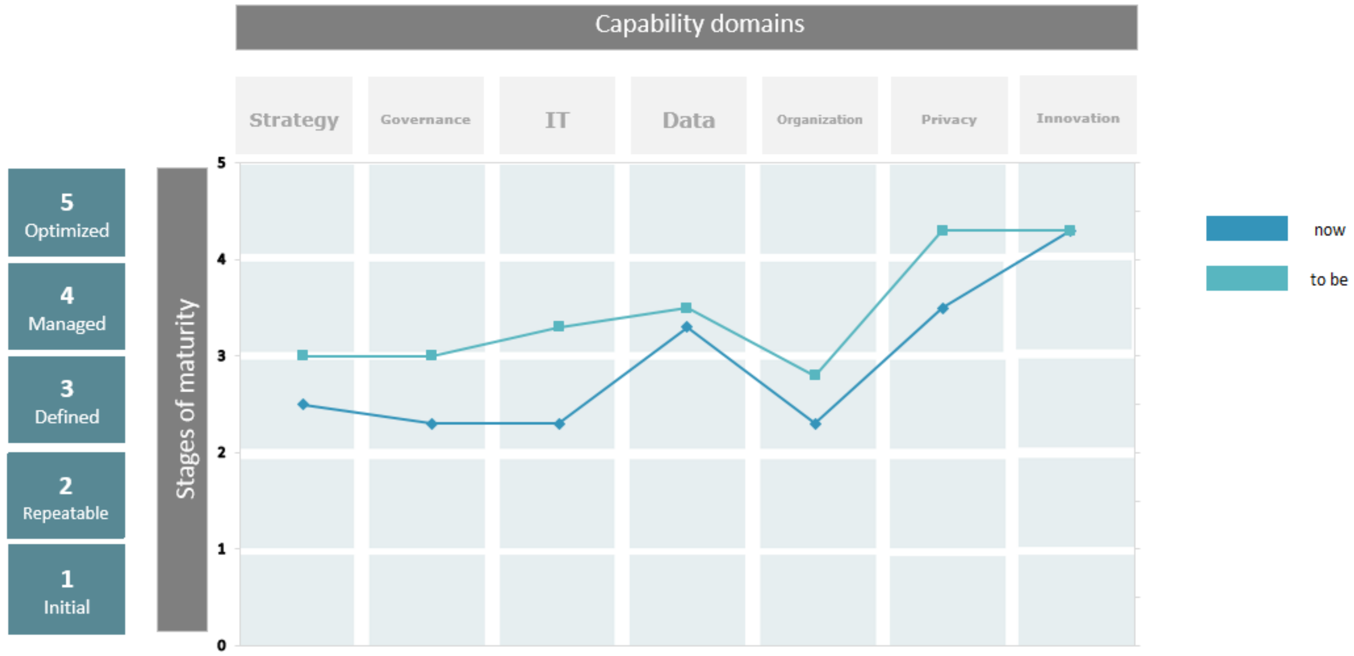


FIGURE 21 BIG DATA MATURITY OF THE RADBOUD UMC

Innovation is very developed, which does not surprise with the presence of the REshape group within the hospital. The REshape group of the Radboud UMC uses technology to change the hospital. They nurture movement by setting up conferences, doing research on different aspects on participatory healthcare and are a vehicle for change. However, the gap in maturity between the REshape department and the rest of the hospital when it comes to having '*an innovative mind'* might be hard to overcome.

Privacy is also on a higher maturity level than the other 5 domains. There is a good discussion on privacy in the academic hospital. There are differences in definitions of privacy and the question 'what is privacy in this digital era?' is raised. Privacy is so much on top of mind that it might make it very difficult for big data analytics to exist within the hospital. Privacy policies on data analytics have to be discussed for the Radboud UMC to become more mature overall.

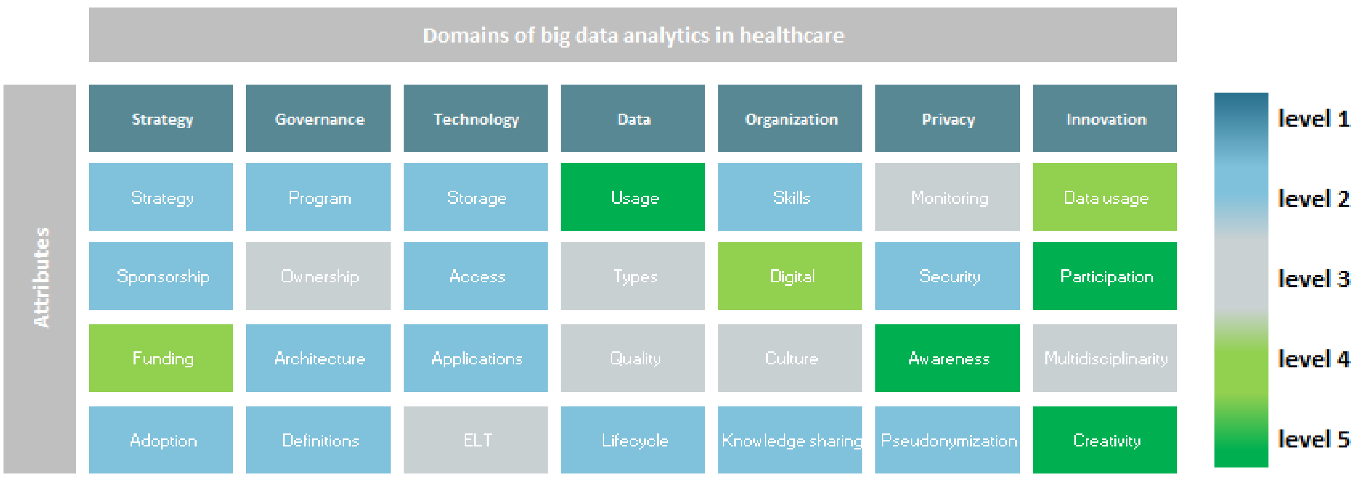


FIGURE 22 CURRENT BIG DATA MATURITY OF THE RADBOUD UMC PER ATTRIBUTE

Figure 22 shows the current maturity of the Radboud UMC on big data analytics per attribute. There is not one attribute that was evaluated as level 1, but there are many attributes on level 2. Figure 23 shows the expected to-be maturity of the Radboud UMC in two years from now.

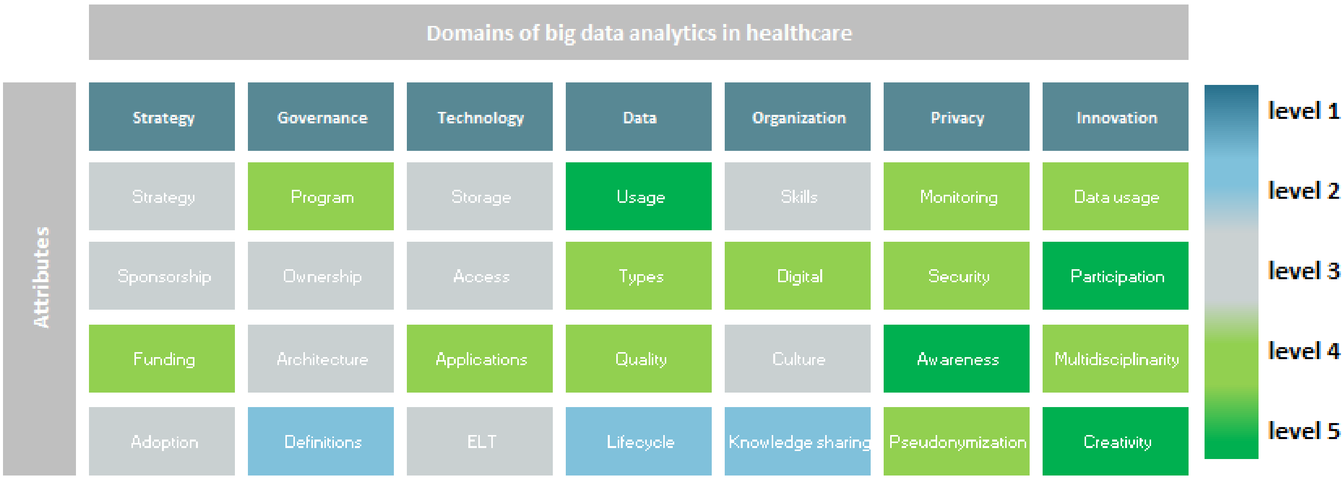


FIGURE 23 TO-BE BIG DATA MATURITY OF THE RADBOUD UMC PER ATTRIBUTE

There are quite a number of changes in some domains, mostly because of projects such as the development of the Digital Research Environment and the project on data governance. The project by the REshape team on big data in the hospital might influence the maturity as well.

Some attributes should not be overlooked such as the lack of knowledge sharing on big data between research groups. The most limiting factor of the researchers seems to be time. As a result of the lack of time, researchers are very focused on their own projects and cannot spend time on things that might not seem as important. A solution to this might be at least a technology center on big data analytics, but if we look even further ahead it should be noted that researchers could be more actively involved in knowledge sharing on a horizontal level throughout the hospital. With the current organizational structure, the UMC will reinvent the wheel many times.

The current organizational structure also leads to other difficulties such as collaboration for projects that are not specific for one department but might need a multidisciplinary team. The question is who will fund these project, because funding comes from departments and this might prove to be a political game. Because the analyses are executed within departments, these might narrow-minded. Multidisciplinary teams may lead to more innovative ideas. This might be accelerated by hiring data scientists that are not committed to one department, but have the position to look broader.

Governance, Technology and Organization are the domains that will mature a lot the coming years due to the current projects that we described previously. However, attributes such as Funding, Multidisciplinarity and Knowledge sharing should not be overlooked.

REFLECTION

All interviewees were sent the results of the maturity assessment. This included the filled in scoring form, documentation of the model and a report with the findings. These documents can be seen in Appendix F. The interviewees were asked to fill in a survey on the results of the maturity assessment. This survey, and the results, can be seen in Appendix D. For each of the questions from the maturity model, the respondents were asked to provide the best-fitting answer according to them. The survey was only answered by the sponsor of the project, as the rest did not feel like an expert on all topics so they thought they were not capable to answer the questions. The response was compared to the chosen answer in the maturity assessment, visible in Table 9. The assessor and the interviewee agreed on the current maturity of the Radboud UMC.

Question	Strategy	Governance	Data	IT	Organization	Privacy	Innovation
1	2 – 2	2 – 2	5 – 5	2 – 2	2 – 2	3 – 3	4 – 4
2	2 – 2	3 – 3	3 – 3	2 – 2	2 – 2	4 – 4	5 – 5
3	4 – 4	2 – 2	3 – 3	2 – 2	3 – 3	5 – 5	3 – 3
4	2 – 2	2 – 2	2 – 2	3 – 3	2 – 2	2 – 2	5 – 5

TABLE 9 ANSWERS TO QUESTIONS OF THE MATURITY MODEL OF RADBOUD UMC FROM ASSESSOR COMPARED TO AVERAGE FROM INTERVIEWEES – DIFFERENCES GREAT THAN 0,3 HIGHLIGHTED IN RED

Similarly, the interviewees were asked to answer the questions on the to-be maturity of the VUMC. Again, only the sponsor of the assessment answered the survey because others considered themselves as not knowledgeable enough on the topic. The response was compared to the chosen answer in the maturity assessment, visible in Table 10.

Question	Strategy	Governance	Data	IT	Organization	Privacy	Innovation
1	3 – 3	3 – 4	5 – 5	3 – 4	3 – 2	4 – 4	4 – 4
2	3 – 3	4 – 3	3 – 4	3 – 5	3 – 3	4 – 4	5 – 5
3	4 – 5	3 – 4	4 – 4	4 – 4	3 – 3	5 – 5	3 – 4
4	2 – 4	2 – 4	2 – 2	3 – 4	2 – 2	4 – 4	5 – 5

TABLE 10 ANSWERS TO QUESTIONS OF THE MATURITY MODEL OF THE RADBOUD UMC ON THE TO-BE MATURITY FROM ASSESSOR COMPARED TO AVERAGE FROM INTERVIEWEES – DIFFERENCES GREAT THAN 0,3 HIGHLIGHTED IN RED

There is less agreement on the to-be maturity of the Radboud UMC than on the current maturity of the Radboud UMC. There are disagreements on the attributes Funding and Processes from the domain Strategy. The Radboud UMC will actively try to have all departments engage with big data analytics.

Within the Governance domain, every attribute is answered with a different maturity level by the sponsor and the assessor. The sponsor values most attributes higher than the assessor. Radboud UMC has contact with the Ministry of Health to have more centralized governance on all projects in the near-future. Disagreements on the Governance domain are mostly due to the fact that there are two central data storages that can be considered for this question, the Digital Research Environment and the Business Intelligence data warehouse.

The attribute Data Types is expected to reach level 4 of maturity in two years' time. The domain Information Technology is expected to reach a higher maturity in three attributes by the assessor because this is a focus area of the Radboud UMC. The sponsor also does not expect to have data scientists in the hospital within two years from now. Knowledge sharing is also expected to happen between multidisciplinary teams.

The answers of the survey were used to adjust the maturity assessment. Then, the model was adjusted to version 2.1.

FINAL BIG DATA MATURITY OF THE RADBOUD UMC

The suggestions provided by the respondents were used to adjust the answers to the questions of the maturity model. Changes were only made when the assessor missed some information because it did not show in the interviews.

Changes were made to:

- The to-be maturity of the attribute Adoption was changed from 2 to 3.

- The to-be maturity of the attribute Big data governance program was changed from 3 to 4.
- The to-be maturity of the attribute Data ownership was changed from 4 to 3.
- The to-be maturity of the attribute Data types was changed from 3 to 4.
- The as-is maturity and the to-be maturity of the attribute Culture were changed from 3 to 2.
- The as-is maturity of the attribute Digital is assessed at maturity level 4 and the to-be maturity of this attribute is also assessed at maturity level 4.
- The as-is maturity of the attribute Security is assessed at maturity level 2 and the to-be maturity of this attribute is assessed at maturity level 4.
- The to-be maturity of the attribute Multidisciplinary teams is updated from 3 to 4

The final current big data maturity level of the Radboud UMC is 2,9 and the final to-be big data maturity level of the Radboud UMC is 3,6.

CASE STUDY AT VUMC

For the case study at VUMC, the maturity model version 1.0 was used. The as is maturity and the to-be maturity of the VUMC can be seen in Figure 24. The overall maturity of the VUMC is 3,1 and the to-be maturity of the VUMC is 3,3.

Some domains are more mature than others. Data and Privacy are much more mature than the other four domains. Both of these domains, Data and Privacy, are level 4. Strategy, Data Governance, IT and Organization are level 2 and Innovation is level 3.

Especially the Data domain is relatively very mature, when compared with for example IT. The Business Intelligence department is mature and well organized. There is only one data warehouse and that is very different compared to other academic hospitals in the Netherlands. There is a clear focus on making all sorts of data available in the warehouse to all potential customers and becoming a data-driven organization.

Privacy is also on a higher maturity level than the other domains. Privacy is so much on top of mind that it might make it very difficult for big data analytics to exist within the hospital. Privacy protocols have extended such that usage of data is traceable, flags go off when necessary and within the business intelligence some computers have closed off USB ports.

The explicit choice to keep data access to the warehouse a manual process that needs the guidance of the BI makes exploring time-consuming, as this will always be a process that takes time. Also, the fact that the BI is really only deals with the data until the researcher does the analysis makes an interface with big data analytics applications a situation that will not be reached. The applications are now really embedded in the departments, with differences in software per department. This could be taken into account in the research data platform plan.

The gap between the current maturity and the to-be maturity of the VUMC is quite small, and at points (Privacy and Innovation) even zero. It is good to realize that the planned effort on the business intelligence in the future will not ensure a much higher maturity. This means that the VUMC should realize that if it wants to become more mature in big data analytics, it should plan to take action.

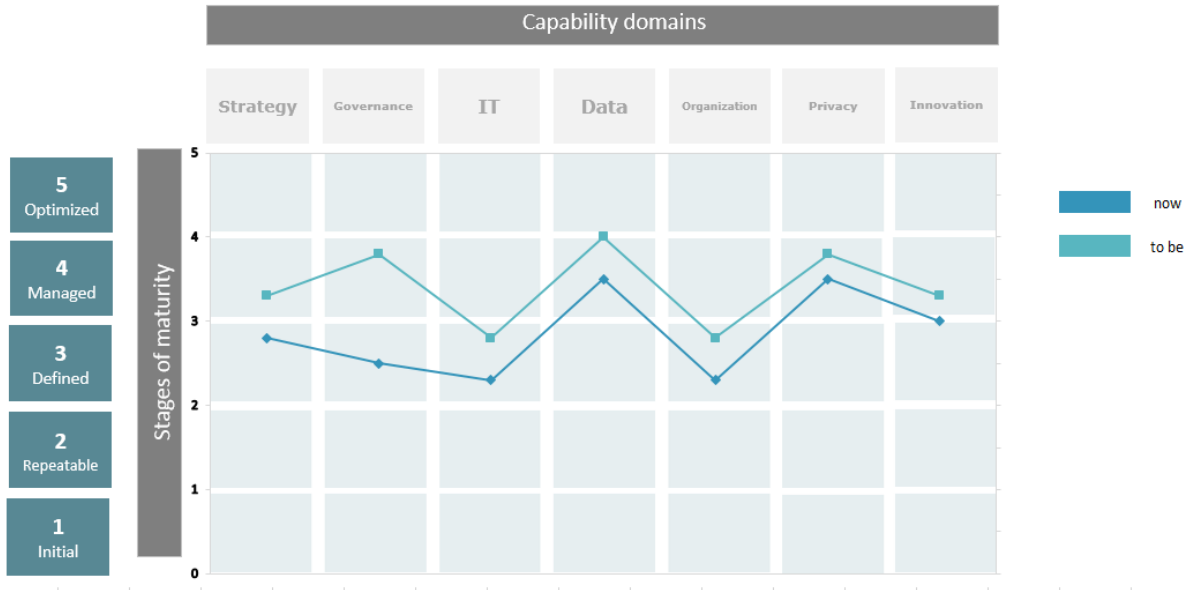


FIGURE 24 BIG DATA MATURITY OF THE VUMC

Figure 25 shows the current maturity of the VUMC on big data analytics per attribute. There is a big difference in maturity per attribute, as there are two attributes that are at level 1 and four at level 5. Figure 26 shows the expected to-be maturity of the VUMC two years from now.

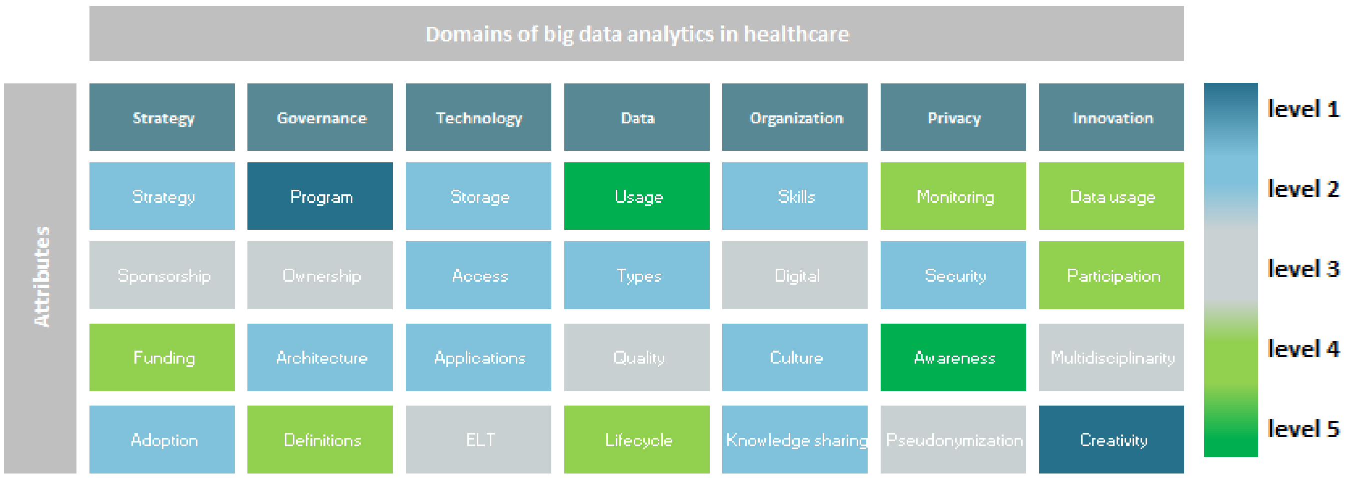


FIGURE 25 CURRENT BIG DATA MATURITY OF THE VUMC PER ATTRIBUTE

The number of changes is quite low, as expected from the overall change in maturity as described in Figure 24. The change in process is due to the fact that the BI department will focus on who is and who isn't using data and provide follow-up consultancy towards customers that have access to data from the data warehouse.

The relatively lower maturity of the Governance domain might complicate the situation in the future, as a lack of policies might lead to misunderstanding in terms of responsibilities or tasks that are not executed. The difference between Technology and Data might also lead to problems as the BI department is not that big and by keeping access to data a manual process, the expected rise in data requests will lead to an overload of requests for the department. There is a clear emphasis on the attributes from the Data domain, which is immediately visible in the maturity of the VUMC, but it should be noticed that Technology and Governance should not be overlooked.

The domains organization and innovation have some domains that are on a low level. This is mostly due to a lack of knowledge sharing in the hospital. Knowledge in the hospital is shared to other hospitals working in the same field of expertise, but it is a pity that this knowledge is not also shared with the other departments of the VUMC. With the current ways of working, the wheel will possibly be reinvented many times. There is a multidisciplinary team in the BI department which could have more responsibilities such as a central data science role to support the researcher performing the data analysis.

Even though there is an innovation budget of 1 million euros per year, it could be useful to have a separate team working on innovation. Literature describes that it is important, especially in big organizations that slowly change such as academic hospitals, to have teams that can work independently of the organization.

If big data is really on the agenda for the years to come, it should be noted that with the current plans the hospital will not mature much on this matter.

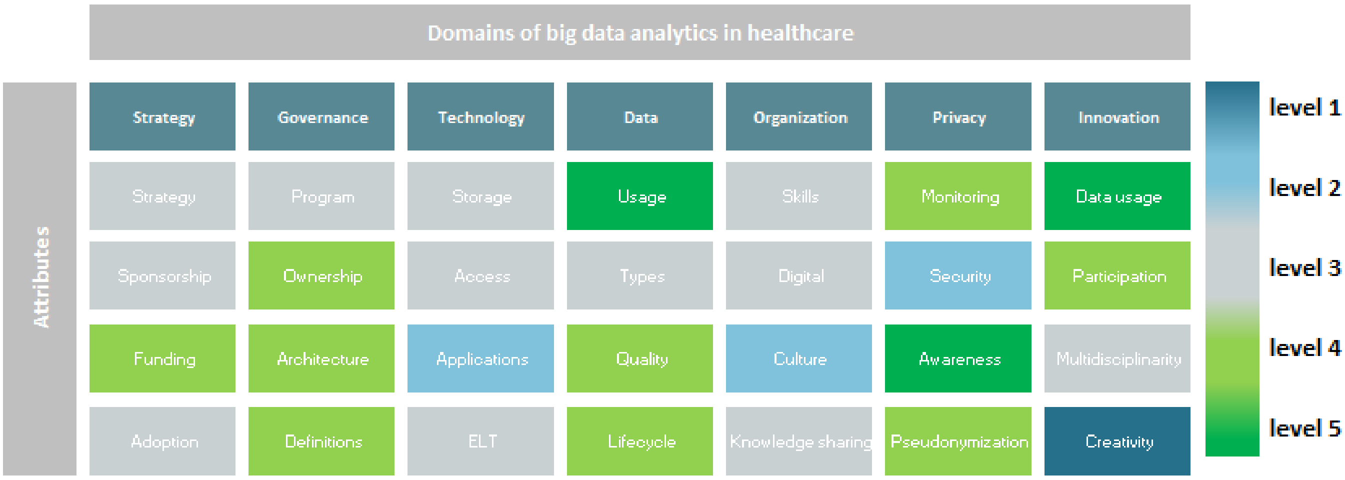


FIGURE 26 TO-BE BIG DATA MATURITY OF THE VUMC PER ATTRIBUTE

REFLECTION

All interviewees were sent the results of the maturity assessment. This included the filled in scoring form, documentation of the model and a report with the findings. These documents can be seen in Appendix F. The interviewees were asked to fill in a survey on the results of the maturity assessment. This survey, and the results, can be seen in Appendix D. For each of the questions from the maturity model, the respondents were asked to provide the best-fitting answer according to them. The average of these responses was compared to the chosen answer in the maturity assessment, visible in Table 11. Mostly, the assessor and the interviewees agreed.

Question	Strategy	Governance	Data	IT	Organization	Privacy	Innovation
1	2 – 2	1 – 1,7	5 – 5	2 – 2,3	3 – 2,3	4 – 4	5 – 4,3
2	3 – 2,7	3 – 3	2 – 2,3	2 – 2	2 – 2	5 – 4,7	4 – 4
3	4 – 4	2 – 2,7	4 – 3,3	2 – 2,3	3 – 3	5 – 5	3 – 3
4	2 – 2,7	4 – 3,7	4 – 3,7	3 – 2,7	2 – 2	4 – 3,3	1 – 2,3

TABLE 11 ANSWERS TO QUESTIONS OF THE MATURITY MODEL OF THE VUMC ON THE CURRENT MATURITY FROM ASSESSOR COMPARED TO AVERAGE FROM INTERVIEWEES – DIFFERENCES GREAT THAN 0,3 HIGHLIGHTED IN RED

There was disagreement on eight attributes. One attribute is from the Strategic Alignment domain, namely Process. Two respondents answered this question with maturity level 3, and one respondents assessed it at maturity level 2 which is in agreement with the assessor.

Two attributes of the Governance domain, namely Data governance program and Data Architecture, were disagreed on. Two respondents agreed with the assessor on the question on Data governance program, and one respondent had a significantly other answer that is two maturity levels higher than all others answered. The question on Data Architecture was answered by two respondents with maturity level 3, and one with maturity level 2 which is in agreement with the assessor.

The question on applications was answered by two respondents with maturity level 3, and one of the respondents agreed with the assessor on maturity level 4. The attribute Skills was valued at maturity level 2 by two respondents, while one agreed with the assessor on maturity level 3.

The question from the domain Privacy on Technology was answered by two respondents with maturity level 4, and the one responsible for this technology answered this question with maturity level 2 which is in agreement with the assessor.

The question on Data usage was agreed on by one respondent, but two respondents answered the question with maturity level 4. The question on Creativity in the Innovation domain was agreed on by two respondents, but the manager of this department disagreed and answered with maturity level 1. The respondent explained that there officially is a department on innovation but they are not yet active.

Similarly, the interviewees were asked to answer the questions on the to-be maturity of the VUMC. The average of these responses was compared to the chosen answer in the maturity assessment, visible in Table 12.

Question	Strategy	Governance	Data	IT	Organization	Privacy	Innovation
1	2 – 3	1 – 3	5 – 5	3 – 3,3	3 – 3,3	4 – 4,3	5 – 4,7
2	3 – 3,7	3 – 3,7	3 – 3,3	2 – 2,7	2 – 2,7	5 – 4,7	4 – 4
3	4 – 4	4 – 4	4 – 4	2 – 2,3	4 – 4	5 – 5	3 – 3,3
4	2 – 3,3	4 – 4	4 – 3,7	3 – 3	2 – 3	4 – 3,7	1 – 2,7

TABLE 12 ANSWERS TO QUESTIONS OF THE MATURITY MODEL OF THE VUMC ON THE TO-BE MATURITY FROM ASSESSOR COMPARED TO AVERAGE FROM INTERVIEWEES – DIFFERENCES GREAT THAN 0,3 HIGHLIGHTED IN RED

There is less agreement on the to-be maturity of the VUMC than on the current maturity of the VUMC. Three out of eight attributes that were disagreed on for the current maturity of the VUMC are still disagreed on, but there are also six other attributes that are disagreed on.

The attribute Strategy was valued by the respondents with 2, 3 and 4. The attribute Governance program was valued with 2, 3 and 4 by the respondents. The attribute Sponsorship was valued by respondents with two 4's and a 3. The attribute Process was valued by respondents with two 3's and a 4, all a higher maturity than the chosen maturity by the assessor.

The attribute Data governance program was responded to with 2,3 and 4 while the assessor valued the attribute with maturity level 1. The attribute Ownership was responded to with two 4's and one 3. The attribute Access was responded to with two 3's and one 2.

The attribute Stewards was responded to with two 3's and one 2. The attribute Knowledge sharing was responded to with all 3's while the assessor answered with maturity level 2.

The Creativity attribute received a 1, a 3 and a 4 while the assessor answered the question with maturity level 4. The 1 was answered by the manager of ICT innovation.

The answers of the survey were used to adjust the maturity assessment. Then, the model was adjusted to version 2.1.

FINAL BIG DATA MATURITY OF THE VUMC

The suggestions provided by the respondents were used to adjust the answers to the questions of the maturity model. Changes were only made when the assessor missed some information because it did not show in the interviews.

Changes were made to:

- The to-be maturity of the attribute Strategy was changed from 2 to 3.
- The to-be maturity of the attribute Adoption was changed from 2 to 3.
- The to-be maturity of the attribute Big data governance program was changed from 1 to 3.
- The current maturity of the attribute Quality was changed from 4 to 3.
- The to-be maturity of the attribute Ownership was changed from 3 to 4.
- The to-be maturity of the attribute Data access was changed from 2 to 3.
- The current maturity of the attribute Skills was changed from 3 to 2.
- The to-be maturity of the attribute Knowledge sharing was changed from 2 to 3.
- The current maturity and the to-be maturity of the attribute Culture was changed from 3 to 2.
- The as-is maturity of the attribute Digital is assessed at maturity level 3 and the to-be maturity of this attribute is also assessed at maturity level 3.
- The as-is maturity of the attribute Pseudonymization is changed from 4 to 3.
- The as-is maturity of the attribute Security is assessed at maturity level 2 and the to-be maturity of this attribute is assessed at maturity level 2.
- The as-is maturity of the Usage attribute is changed from 5 to 4.
- The to-be maturity of the attribute Multidisciplinary teams is updated from 3 to 4

The final current big data maturity level of the VUMC is 2,8 and the final to-be big data maturity level of the VUMC is 3,4.

CASE STUDY AT LUMC

For the case study at LUMC, the maturity model version 2.0 was used. The as is maturity and the to-be maturity of the LUMC can be seen in Figure 27. The overall maturity of the LUMC is 2,7 and the to-be maturity of the LUMC is 2,9.

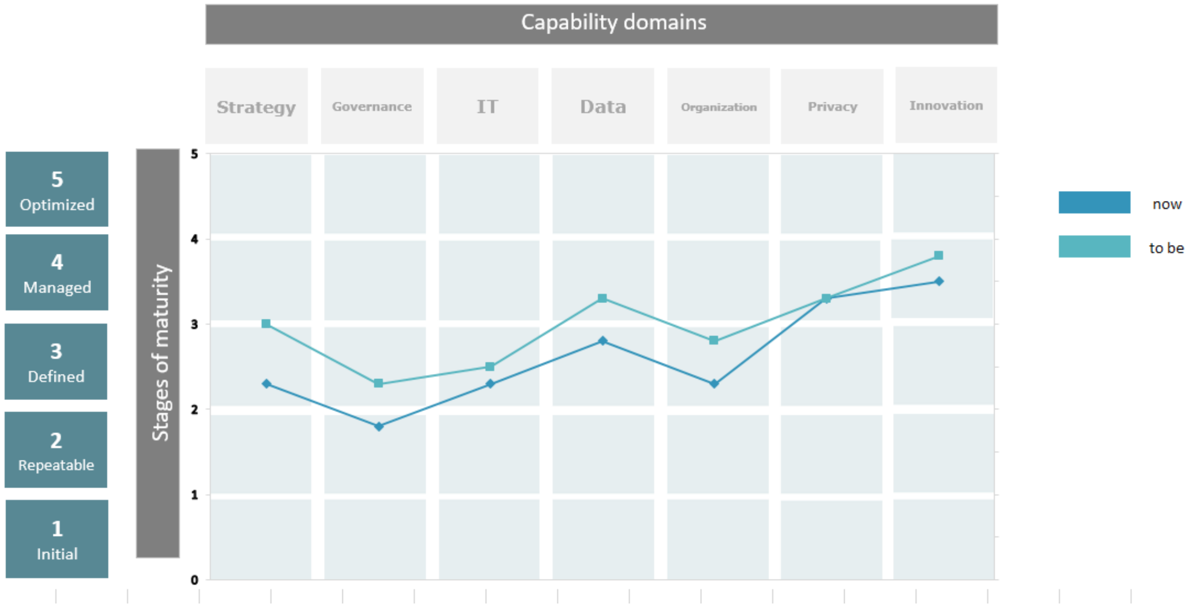


FIGURE 27 BIG DATA MATURITY OF THE LUMC

Some domains are more mature than others. Innovation and privacy are much more mature than the five other domains. The other five domains are all on level 2.

Innovation is relatively developed, which does not surprise with the presence of the newly shaped team on innovation within the IT department. However, as this team is still in the start-up phase, it might prove to be difficult to not get sucked into the daily operations of the hospital that is often referred to as the 'cruise ship'. Key is to keep behaving as a speed-boat next to the cruise ship. It might also be challenging to get buy-in from the rest of the hospital when it comes to the innovative mind. The innovation team might be able to stimulate knowledge-sharing and the use of multidisciplinary teams in other big data analyses.

Privacy is also on a higher maturity level than the other domains. Privacy is so much on top of mind that it might make it very difficult for big data analytics to exist within the hospital. Really exploring data to find new patterns will find resistance because of privacy concerns. The pseudonymization techniques that are currently used by the BI department should be addressed to take away some of this concerns. Also, automatic pseudonymization of free text might be a solution for this issue. The data that is extracted from the central data warehouse by the MI group is now handed over. The data can then be used in any way, in an uncontrolled environment. This needs to be addressed by the MI group, because with the current plans for the data warehouse there might be a rise in the number of data requests. Manually tracking the data will no longer be sufficient. Also, the rise in the number of requests for data will be a burden on the MI group. Manual extraction of data will not be sufficient.

The current situation with two teams (MI and MIS) that should work together, but which proved to be difficult in reality, will split up in the near future. The LUMC should realize that having two warehouses, with two sets of people, protocols and processes might not be a desired situation. This is also a big contrast with the current program 'Registratie aan de bron', where data can be recycled after registering it correctly the first time. If two teams will separately work on data extraction and quality, there can be multiple versions of the same original data.

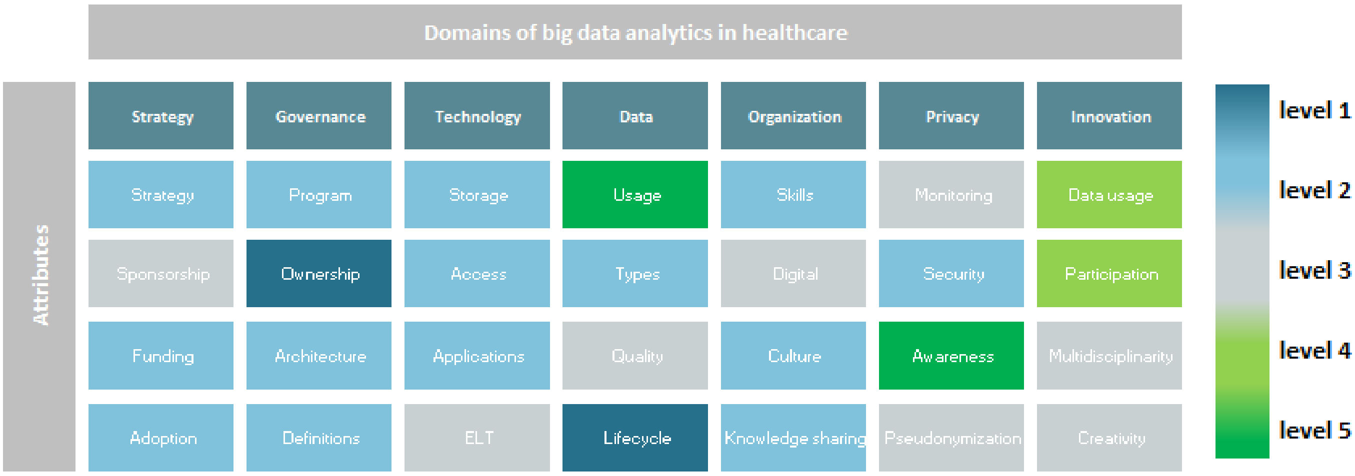


FIGURE 28 CURRENT MATURITY OF THE LUMC PER ATTRIBUTE

Figure 28 shows the current maturity of the LUMC on big data analytics per attribute. There is a substantial difference in maturity per attribute, as there are two attributes that are at level 1 and two at level 5. Figure 28 shows the expected to-be maturity of the LUMC two years from now.

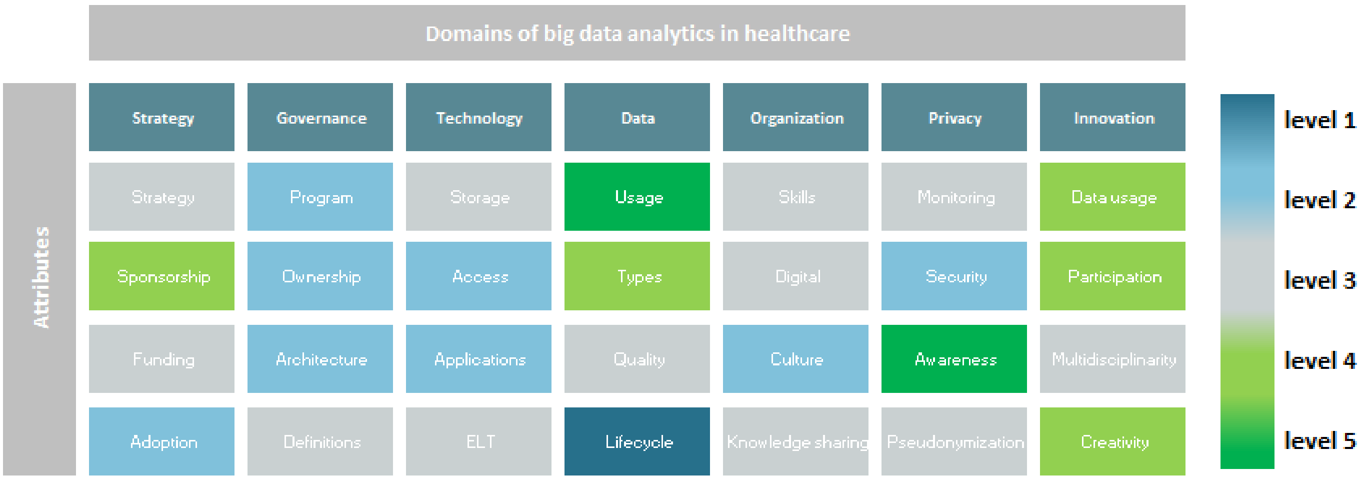


FIGURE 29 TO-BE MATURITY OF THE LUMC PER ATTRIBUTE

Figure 28 shows that the LUMC will not progress in all areas. A higher maturity is only reached for Strategic alignment, Data and Innovation. This is remarkable because big data analytics is mentioned in the strategic document of the LUMC on the public homepage, but this focus is not recognized by the maturity model.

There are some attributes that should not be overlooked in the next two years. The Security attribute of the Privacy domain is important to mention. This is very much related to the Lifecycle attribute of the Data domain, which only has a maturity level of 1. If there is a controlled environment where data can be used for analytics purposes, the lifecycle of this data can also be maintained.

The attributes that should be considered by the LUMC are from the Organization domain. Speaking with non-IT employees in the hospital about big data often leads to questions on the usefulness. These medical specialists are trained to think about data analysis in a top-down way, using the data to prove a hypothesis with statistics. The mindset of these employees needs to be addressed for big data analytics to truly be accepted by the LUMC.

These employees with doubts on big data analytics are mostly medical specialists and not researchers. Researchers are using big data analytics, but do not share their ideas or knowledge with the rest of the LUMC. You could say that there is a sort of gap between those worlds, with different mindsets: one focused on treating patients right now and the other focused on finding treatments that will help patients' years from now. The knowledge already obtained by these research groups could be distributed to other departments.

The difference between the Data and the Information Technology department can be troublesome. If the data from the data warehouse is more heavily used and more data types are available, the Information Technology domain should not be forgotten. As previously described, the burden on the MI team will increase if this domain is not addressed.

With the current plans on big data analytics, the LUMC will not reach a higher maturity in the next two years. This means that big data analytics can be done in small projects such as proof-of-concepts, but will not be a standard process for most departments.

REFLECTION

All interviewees were sent the results of the maturity assessment. This included the filled in scoring form, documentation of the model and a report with the findings. These documents can be seen in Appendix F. The interviewees were asked to fill in a survey on the results of the maturity assessment. This survey, and the results, can be seen in Appendix D. For each of the questions from the maturity model, the respondents were asked to provide the best-fitting answer according to them. The average of these responses was compared to the chosen answer in the maturity assessment, visible in Table 13. Mostly, the assessor and the interviewees agreed.

Question	Strategy	Governance	Data	IT	Organization	Privacy	Innovation
1	2 – 2	2 – 2	5 – 5	2 – 2,5	2 – 2,5	3 – 3	4 – 4
2	3 – 3	1 – 1	2 – 2,5	2 – 2	3 – 3	2 – 2	4 – 4,5
3	2 – 2	2 – 2,5	3 – 3	2 – 2,5	2 – 2	5 – 5	3 – 2,5
4	2 – 2	4 – 2	1 – 2	3 – 2,5	2 – 2	3 – 3	3 – 3

TABLE 13 ANSWERS TO QUESTIONS OF THE MATURITY MODEL OF THE LUMC FROM ASSESSOR COMPARED TO AVERAGE FROM INTERVIEWEES – DIFFERENCES GREAT THAN 0,5 HIGHLIGHTED IN RED

There was disagreement on two attributes. One attribute of the Data domain, namely Data lifecycle. One respondent agreed with the assessor on this topic, while one correspondent valued the question with maturity level 3. The respondent did not elaborate on this answer. The other attribute that the assessor and the respondents disagreed on significantly is the attribute Data definitions from the Governance domain. Both respondents valued the attribute to be at maturity level 2. This was elaborated by a respondent with "This was a difficult question to answer. We have many standards, but it almost impossible to cover all data with these definitions."

Similarly, the interviewees were asked to answer the questions on the to-be maturity of the VUMC. The average of these responses was compared to the chosen answer in the maturity assessment, visible in Table 14.

Question	Strategy	Governance	Data	IT	Organization	Privacy	Innovation
1	3 – 3	2 – 2	5 – 5	2 – 3,5	2 – 3	3 – 3	4 – 4,5
2	4 – 4	2 – 2,5	4 – 4	2– 2,5	3 – 3	2 – 2,5	4 – 4,5
3	3 – 3	2 – 2	3 – 3	2 – 2,5	2 – 2,5	5 – 5	3 – 3,5
4	2 – 2	3 – 3	1 – 2	3– 3	2 – 3	3 – 3	4 – 4,5

TABLE 14 ANSWERS TO QUESTIONS OF THE MATURITY MODEL OF THE LUMC ON THE TO-BE MATURITY FROM ASSESSOR COMPARED TO AVERAGE FROM INTERVIEWEES – DIFFERENCES GREAT THAN 0,5 HIGHLIGHTED IN RED

There is less agreement on the to-be maturity of the LUMC than on the current maturity of the LUMC. There is disagreement on 4 attributes, of which one was also disagreed on for the current maturity of the LUMC. The attribute Data lifecycle is disagreed on, just like for the current maturity of the LUMC. However, respondents and assessor agree that the maturity of the LUMC will stay the same on this matter. They only disagree on what level that is.

The attribute Data storage from the domain Information Technology was valued at maturity level 2 by the assessor, while the respondents valued this attribute at maturity level 3 and 4. This answer was not elaborated on by the respondents.

Within the Organization domain, two attributes were disagreed on. These attributes are the attribute Skills and the attribute Knowledge sharing. Both these attributes were valued by the respondents with maturity level 3, while they were valued by the assessor with maturity level 2. These answers were not elaborated on.

The answers of the survey were used to adjust the maturity assessment. Then, the model was adjusted to version 2.1.

FINAL BIG DATA MATURITY OF THE LUMC

The suggestions provided by the respondents were used to adjust the answers to the questions of the maturity model. Changes were only made when the assessor missed some information because it did not show in the interviews.

Changes were made to:

- The current maturity of the attribute Data definitions was changed from 4 to 2.
- The to-be maturity of the attribute Data storage was changed from 2 to 3.
- The to-be maturity of the attribute Skills was changed from 2 to 3.
- The to-be maturity of the attribute Knowledge sharing was changed from 2 to 3.

The final current big data maturity level of the LUMC is 2,6 and the final to-be big data maturity level of the LUMC is 3.

CONCLUSIONS

The developed maturity model was used to assess three Dutch academic hospitals. These UMCs had maturity of 2.6, 2,8 and 2,9. All these levels of maturity are closest to maturity level 3 ‘defined’. At this maturity level, big data analytics can be performed within the hospital and there should be a standard business process. There are processes in place to facilitate big data analytics throughout the hospital, but these still require manual labor.

Two of the three academic hospitals examined have two ‘data warehouses’, one for medical data and one for management information. Furthermore, big data analytics is definitely not yet on the radar for all employees of the academic hospitals. There is still a very conservative attitude towards big data and statistical analysis are preferred over explorative analysis methods.

Radboud UMC will make the most progression in big data analytics maturity in the next two years, as they will go from 2,9 to 3,6. VUMC will go from 2,8 to 3,4 and LUMC from 2,6 to 3,0. VUMC is leading in the Privacy and Data domain, Radboud UMC is leading in the Innovation domain. Information

Technology, fundamental for big data analytics, is remarkably of relatively low maturity in all three academic hospitals.

Reaching a higher maturity and realizing changes in academic hospitals mostly takes more than two years' time. It seems time-consuming to change a massive organization, especially with so many researchers that have personal motives to publish and attract funds and doctors that have an extremely high workload and almost no spare hours in a week for innovation.

6.3 NATIONAL INITIATIVES

This section discusses the national initiatives on big data analytics in healthcare. During this research project, we came across many national initiatives that are in some way working on big data analytics with UMCs or other healthcare providers. These initiatives mostly work on making data available for other healthcare professionals. Not one of the initiatives is an initiative explicitly for big data analytics, but these initiatives are working on the preconditions of a high maturity of big data analytics in academic hospitals in the Netherlands. The discussed information was obtained by interviews with chairmen from these national initiatives, see Appendix C.

Direct collaboration

One of the oldest and highly respected initiatives, in which all UMCs participate, is the Nederlandse Federatie van Universitair Medische Centra (NFU). The NFU has multiple meetings with attendees from different layers of the organization. For example: the CEO's come together, the CFO's have meetings and, important for this research project: the CIO's have a monthly meeting, called the AcZie. Sometimes IT projects can be done together, such as buying a new Electronic Patient Record system or hardware for servers. However, these collaborations can only happen when UMCs are coincidently thinking about renewing that particular thing at the same moment in time. It can be extremely difficult to collaborate with all UMCs in one project as every UMC is an individual with own wishes and requirements, and they are able to step out of a project if they wish. Nothing that is agreed upon within the NFU is binding, as all parties are equal. Because there are too many topics to discuss and the CIO's are not always as deep in the topics as needed, there are some other IT meetings that occur a couple of times a year. Examples of these meetings are the TacZie, which is similar to the AcZie but then discussing IT on a technical level and not on a strategic level, or Sig Prima, a meeting on ICT architecture.

There is a big difference in structure of the AcZie and the other meetings on IT subjects. The AcZie is always with the same eight persons: the eight CIO's from UMC's in the Netherlands. However, for the other meetings on IT subjects the attendees vary. Depending on the topic of the meeting, different people from the UMCs attend. As a result of many topics to discuss, one meeting has an agenda with another agenda item every ten minutes. This way, lots of topics are discussed briefly and because of the huge variety of topics sometimes over 60 people are invited for the meetings. Interviewees respond to the usefulness of these meetings with *"It is mostly a lot of complimenting each other and not so much doing things. Mostly a lot of chitchat without ending up with a clear plan or procedure"*. One of the UMCs is now making plans to make these meetings more productive.

The medical intelligence teams of the academic hospitals, which are responsible for a central data storage of medical information, also meet regularly. This meeting is not under the responsibility of the NFU, but arranged by a commercial party.

NFU initiated collaborations

The NFU has initiated some national programs, of which some are in some way working on big data analytics. One of these programs is data4lifesciences. Data4lifesciences aims to realize an integrated research data infrastructure in, for, by and between UMCs and their partners. They collaborate with national initiatives on research data.

Another program is registration at the source, which focusses on doing administrations one time properly to use this data multiple times for other purposes. They are creating standards, implement those in the IT systems and change the ways of working so that information only needs to be registered once.

Thirdly, there is a program on e-health focusing on providing access to healthcare data at any time for all patients. The UMCs aspire to create one digital personal health environment to enable the patient to manage their disease and health.

These programs are all in progress have and not yet been realized.

National initiatives on research data
Nationally there are numerous initiatives working on the availability of research data. There are national databases such as Parelsnoer and BBMRI-NL that gather data from UMCs and other healthcare providers. Furthermore, there are initiatives working on research infrastructure. Examples are an infrastructure for life science research (DTL) and one for translational research IT infrastructure (TraIT). Other initiatives are the personal health train, an initiative that is now part of the data4lifescience program from the NFU, which aims to bring the research to the data and not the other way around. The 'train' should visit the different stations (databases with the requested data). Data should be FAIR: findable, accessible, interoperable and readable for the personal health train to work.

In 2020 it is decided that everyone should have access to all research data. This is called the 'Nationaal plan Open Science'. The European Committee has provided 2 billion euros to realize open science in Europe. This responsibility lies with the initiative Health-RI. The mission of Health-RI is to enable excellent personalized medicine & health research by facilitating the research process from start to end. To this end Health-RI acts as a public utility with a portfolio of services including catalogues of data, images and samples and a digital research environment. Health-RI is an initiative by many different founders such as BBMRI, DTL, the NFU and many others.

National initiatives on healthcare data
There are many national initiatives that contribute to the ability to share healthcare data with other care providers. These are not specifically for big data analytics, but to support healthcare providers when treating a patient. Many of these initiatives are collaborating on different data or different parts of data sharing. For example, there is Zorgdomein that can be used to refer patients to other care providers, there are agreements on using XDS, a standard format for healthcare data, there is the LSP which is an infrastructure to share patient records on between general practitioners. The Ministry of Health has initiated the 'Informatieberaad' where healthcare providers can participate. Participation is open to anyone, but once entered the agreements must be honored. They initiate programs such as MedMij: providing guidelines that need to be fulfilled in order to be able to exchange healthcare data between patients and healthcare providers.

Observations
There are many different national initiatives working on some part of big data analytics. There are initiatives on national databases, infrastructure, standards and knowledge sharing for research data as well as healthcare data. As a result, there are many different ideas on these topics adopted by a varying subsets of healthcare professionals or organizations. Most initiatives are not progressing

fast, it takes years to make a change to the industry when multiple healthcare providers are involved.

There are no initiatives directly combining these two ‘worlds’: research and healthcare data. The national initiatives for healthcare data have a very different character than those for research data. Initiatives that want to work on sharing healthcare data are more pragmatic, but can be quick-and-dirty so solutions from the past are not always sufficient now and have to be changed. Initiatives that want to work on sharing research data are more talk than action: a lot of parties are involved and there are many ideas, but not so many concrete results.

6.4 FUTURE OF BIG DATA IN HEALTHCARE

During this research project, some ideas on the future of big data in healthcare were found in literature or mentioned during interviews. Ideas on this topic were also shared during the scheduled sounding board group sessions, as discussed in Chapter 2. These visions will be discussed in this section.

Lucien Engelen (director of the REshape department at Radboud UMC), who positions himself at the forefront of innovation in healthcare, argues that UMCs are *egosystems* and not ecosystems. “They only think about themselves and find themselves very important, but do not think enough about the patient. The academic hospitals think they should be the ones creating one infrastructure and only talk about sharing their data, but the number of patients for UMCs is (and should be) decreasing.” The government wants that the third line of healthcare, academic hospitals, will only be used in very complicated cases of care and that more attention goes to preventing diseases so the expectation is that the number of patients in academic hospitals will decrease over the years. “It is not only arrogant, but also stupid to only focus on our own data. As an academic hospital you only have data on patients from the one day in the year they visit you. The personal health environments from Google and Apple will soon have more data from those patients than we will ever have. Medical specialists and researchers are trained to cure patients and do medical research, they are not trained to make decisions on IT infrastructures and they should not want to make those decisions. The world around us is changing ten times as fast than we are, and we should learn from the best and not try to develop everything ourselves.”

Ottes (Ottes, 2016) agrees on this, as he mentions the rise of big data developments for consumers via frameworks such as the ResearchKit from Apple. Patients should, by law, be able to view all their personal healthcare data. He argues that there will be a commercial influence to data and the ways data can be shared. Every person has money, but almost no one will leave all this money at home. It is accepted to choose one of the big banks to take care of your money. In return, they can do investments with this. The same could happen for healthcare data. Everyone will eventually have this data, but not everyone will want to keep all responsibilities on this data but use one of the big personal healthcare environments to take care of their data.

During the sounding board group session’s similar ideas were shared. The personal health environment should be seen as an opportunity for research, and not only something that has to be realized by the healthcare industry. Because research data and healthcare data is now discussed in very separate initiatives, the overlap between those might be missed. In essence, both data is the same with one difference: anonymized or not. The initiatives on sharing research data should not only focus on the data we already have as UMCs, but also this newly generated patient data. Also, as long as they are not focusing on IT but on the things they are a specialist in, namely the healthcare industry, they could move a lot faster. By learning from developments made by data specialists such as Google and Facebook, the healthcare industry can transform itself. It should focus on setting guidelines on how to get access to this data and prepare themselves for these developments.

7. CONCLUSIONS

In the previous chapters we illustrated what big data analytics is within academic hospitals, a way to measure how mature an academic hospital is in big data analytics and we tested this maturity method. Furthermore we discussed the national initiatives supporting some part of big data analytics and the future of IT in healthcare.

The main research question of this research project was to find out what the status quo was of big data analytics in Dutch academic hospitals. We conclude this research by answering the research questions that altogether answer the main research question and by suggesting future work.

7.1 WHAT ARE CHARACTERISTICS OF BIG DATA ANALYTICS FOR THE DUTCH ACADEMIC HOSPITALS THAT A MATURITY MODEL MUST CAPTURE?

The first research question was 'What are characteristics of big data analytics for the Dutch academic hospital that a maturity model must capture?'

Based on the literature review on big data analytics in Dutch academic hospitals and the expert interviews, the following characteristics that should specifically be addressed in a big data analytics maturity model for Dutch academic hospitals are as follows:

- Standardization
- Data access
- Privacy and security
- Data-driven use of big data analytics
- Buy-in from the business
- Technology

A big data analytics maturity model for Dutch academic hospitals should specifically capture these characteristics.

7.2 IS THERE A BIG DATA MATURITY MODEL THAT MEETS THE CHARACTERISTICS SPECIFIC FOR THE DUTCH HEALTHCARE INDUSTRY?

The second research question was "Is there a big data analytics maturity model that meets the characteristics specific for the Dutch healthcare industry?"

Eight maturity models were assessed. None of these models were specifically made for the Dutch healthcare industry. The assessed models were from the CMMI, the most widely accepted maturity model, big data analytics maturity models and business intelligence maturity models. These models were assessed on their scientific design, domains, attributes, ways of maturing and scoring methods. Finally, the models were checked on meeting the characteristics of big data analytics in Dutch academic hospitals.

The assessed maturity models did not meet all requirements and characteristics, so a new big data analytics maturity model specific for the Dutch academic hospitals was developed, using the maturity model from Commuzi et al. as a basis.

This developed maturity model is specific for Dutch academic hospitals to assess their current maturity on big data analytics and their to-be maturity on big data analytics in two years from now. The model has five maturity levels that are based on the CMMI. These maturity levels are:

- Initial
- Repeatable

- Defined
- Managed
- Optimized

The model has seven domains, which are Strategic Alignment, Governance, Information Technology, Data, Organization, Privacy and Innovation. These domains each consist of four attributes. For each of these 28 attributes, there exists one question with five answers corresponding each with a different level of maturity that is the obtained score for that attribute. The score per domain is the division of the summed scores of the four attributes of that domain by the maximum score per domain, 20. The model captures the requirements and characteristics specific for the industry.

The method includes instructions for an assessment, a manual for answering questions, scoring forms and a validation survey.

7.3 HOW MATURE IS BIG DATA ANALYTICS IN DUTCH ACADEMIC HOSPITALS CURRENTLY?

The third research question was “How mature is big data analytics in Dutch academic hospitals currently?”

The developed maturity model was used to assess three Dutch academic hospitals. These UMCs had maturity of 2.6, 2,8 and 2,9. All these levels of maturity are closest to maturity level 3 ‘defined’. At this maturity level, big data analytics can be performed within the hospital and there should be a standard business process. There are processes in place to facilitate big data analytics throughout the hospital, but these still require manual labor.

Two out of three academic hospitals have two ‘data warehouses’, one for medical data and one for management information. Furthermore, big data analytics is definitely not yet on the radar for all employees of the academic hospitals. There is still a very conservative attitude towards big data and statistical analyses is preferred over explorative analysis methods.

Two years from now, these maturities will be 3.0, 3.4 and 3.6. The academic hospitals are focusing on different domains such as Innovation or Privacy and Data. The domains Information Technology is remarkably of relatively lower maturity in the assessed academic hospitals.

It will take more than two years to reach a higher level of maturity as change in these massive organizations takes time. The highest level of maturity will most likely not be reached anytime soon because the hospitals do not use the same standards or infrastructure. Furthermore, terms as data lakes or distributed data storages are not even thought of by the assessed academic hospitals. The added value of big data analytics is not considered important enough by most employees in these academic hospitals.

7.4 HOW DO THE DUTCH ACADEMIC HOSPITALS AND THE NATIONAL INITIATIVES ON BIG DATA ANALYTICS IN HEALTHCARE RELATE AND HOW CAN THEY REINFORCE EACH OTHER?

The fourth research question was “How do the Dutch academic hospitals and the national initiatives on big data analytics in healthcare relate and how can they reinforce each other?”

The Dutch academic hospitals are now collaborating with other academic hospitals and participating in many different national initiatives that contribute to preconditions for a high maturity of big data analytics in healthcare. As a result, there are many different standards, infrastructures and parties working on enabling healthcare data sharing. The academic hospitals and the national initiatives have many different meetings every year. However, these are not attended by the same group of people every time and productivity is lacking. The main issue is that the healthcare industry is not centrally managed, for example by the government. As a result, there are too many different projects happening at the same time resulting in a fragmented healthcare industry.

Furthermore, the healthcare industry is divided in two worlds: one focusing on patient data sharing to provide better healthcare, and the other focusing on research data. Two big initiatives focusing on one of these aspects are Health-RI and Informatieberaad. They do try to choose the same software, but are still operating independent of each other.

While the commercial parties are on the rise when it comes to having healthcare data, with the development of personal health environments by i.e. Apple and Google, and the number of patients in academic hospitals decreasing, the healthcare industry should use these changes as an opportunity.

Instead of only providing patients with their healthcare data as agreed on by law, the academic hospitals could also use this as an opportunity to get data not only to these personal health environments but also from these personal health environments. Instead of only focusing on how to share the data the academic hospitals already have, there should also be a focus on how to get access to this huge amount of patient data. Technology companies are already asking users for their data, with success. Academic hospitals should prepare for the future and focus on how they can turn this threat into an opportunity.

7.5 FUTURE RESEARCH

While this research contributed to finding out what the status quo is for big data analytics in academic hospitals, we provide some recommendations for future research.

The model
The developed big data analytics maturity method was tested at three out of eight academic hospitals in the Netherlands. The other five academic hospitals should also be assessed to answer the research question completely.

Industry
The model was created specifically for academic hospitals in the Netherlands, but it has not been tested at peripheral hospitals for validation. It could also be used to assess an organization in a different, complex industry.

Role of the government
Several times during this research project, the role of the government was mentioned. The government was not actively involved in steering the healthcare industry until recently. With the ‘Informatieberaad’, this has changed for the healthcare data. However, there is not yet such a collaboration for research data where the government is involved. The role of the government on big data analytics could be a subject for further research. This could involve the way the government should be involved with the industry.

Personal healthcare environment
The rise of the personal healthcare environment may disrupt the healthcare industry as the data will no longer be stored at the hospitals, but with commercial parties. This development, and the consequences for the healthcare industry, could be researched in a future research project.

8. REFERENCES

Becker, J., Knackestedt, R., & Pöppelbuss, J. (2009). Developing maturity models for IT management. *Business & Information Systems Engineering*, 213-222.

Belle, A., & Thiagarajan, R. (2015). Big data analytics in healthcare. *BioMed Research International*, 1-16.

Bollier, D. (2010). *The promise and peril of big data.* Washington, DC: The Aspen Institute.

Brooks, P., El-Gayar, O., & Sarnikar, S. (2015). A framework for developing a domain specific business intelligence maturity model: Application to healthcare. *International Journal of Information Management*, 337-345.

Burton, B. (2007). *Results of Business Intelligence and Performance Management Maturity Survey*. Retrieved from Gartner: http://www.gartner.com

Chiong Meza, C., Van Steen, J., & de Jonge, J. (2014). *De Nederlandse Universitait Medische centra.* Den Haag: Rathenau Instituut.

Commuzi, M., & Patel, A. (2016). How organizations leverage Big Data: a maturity model. *Industrial Management & Data Systems*, 1468-1492. Retrieved from http://www.emeraldinsight.com/doi/pdfplus/10.1108/IMDS-12-2015-0495

Cottle, M., Hoover, W., Kanwal, S., M, K., Strome, T., & Treister, N. (2013). *Transforming health care through big data strategies for leveraging big data in the health care industry.* Institute for Health Technology Transformation.

CPCT. (2017). Retrieved from CPCT: http://www.cpct.nl/over-cpct/

Data Management Maturity. (2017). Retrieved from CMMI Institute: http://cmmiinstitute.com/data-management-maturity

de Bruin, T., Rosemann, M., Freeze, R., & Kulkarni, U. (2005). Understanding the Main Phases of Developing a Maturity Assessment Model. *16th Australasian Conference on Information Systems* (pp. 2-11). Sydney: Qeensland University of Technology.

de Rijk, T. (2016, March 18). *Student win StrICTly for Business*. Retrieved from Universiteit Leiden: https://www.universiteitleiden.nl/en/news/2016/03/students-win-strictly-for-business

Deloitte. (2015). *Learn and share strategy: About Healthcare.* Amsterdam: Deloitte.

Dhanuka, V. (2016). *Hortonworks big data maturity model.* Santa Clara: Hortonworks. Retrieved from http://hortonworks.com/wp-content/uploads/2016/04/Hortonworks-Big-Data-Maturity-Assessment.pdf

Douglas, L. (2001). *3d data management: Controlling data volume, velocity and variety.* Gartner.

E, G., & al., e. (2012). Perspectives on Big Data and Big Data Analytics. *Database Systems Journal*, 3-13.

El-Darwiche, B., Meer, D., Koch, V., & Thome, W. (2016). *Big data maturity: An action plan for policymakers and executives.* PWC Strategy.

ETL vs ELT: The difference is in the how. (2016, November 6). Retrieved from Panoply: http://blog.panoply.io/etl-vs-elt-the-difference-is-in-the-how

Fay, D., Borrill, C., Amir, Z., Haward, R., & West, M. (2006). Getting the most out of multidisciplinary teams: A multi-sample study of team innovation in health care. *Journal of Occupational and Organizational Psychology*, 553-569.

Feldman, B., Martin, E., & Skotnes, T. (2012). *Big data in healthcare: hype and hope.* Los Angeles: Business development for digital health.

Grimes. (2008, August 1). *Unstructured Data and the 80 Percent Rule.* Retrieved from Breakthrough Analysis: http://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/

Groves, P., Kayyali, B., Knott, D., & Van Kuiken, S. (2013, 2). The 'big data' revolution in healthcare. *McKinsey Quarterly*.

Harvard Business Review. (2012). *The evelution of decision making: how leading organizations are adopting a data-driven culture.* Harvard Business School Publishing.

Heritage Health Prize. (2017). Retrieved from Heritage Health Prize: http://www.heritagehealthprize.com/c/hhp

IBM. (2015, January 25). *The Four V's of Big Data*. Retrieved from IBM big data hub: http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

IBM. (n.d.). *What is big data?* Retrieved from IBM: https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html

InfoTech. (2013, September 25). *Big data maturiy assessment tool*. Retrieved from InfoTech: https://www.infotech.com/research/ss/leverage-big-data-by-starting-small/it-big-data-maturity-assessment-tool#

Koumpouros, Y. (2014). Big data in healthcare. In Y. Koumpouros, *Healthcare Administration: Concepts, Methodologies, Tools, and Applications.*

Lessanibahri, S., Gastaldi, L., Pietrosi, A., & Corso, M. (n.d.). *Co-developing a roadmap towards precision medicine: measuring the maturity of BI in healthcare.*

Magee, F. (2016). *IDC MaturityScapes: A practical guide to deriving value.* IDC.

Mettler, T. (2009). *A Design Science Research Perspective on Maturity Models in Information Systems.* Gallen: Institute of Information Management.

Nguyen, C., & Mendes, M. (2011). *Thiopurine methyltransferase (TPMT) genotyping to predict meylossupression ris.* PLoS currents.

Nu.nl. (2017). Elektronisch Patiëntendossier gewoon lek. Nu.nl. Retrieved from http://www.nu.nl/internet/2213458/elektronisch-patientendossier-gewoon-lek.html

Ottenheijm, S. (2015). *Big data in de gezondheidszorg.* Den Haag: Nictiz. Retrieved from Nictiz: http://www.nictiz.nl/SiteCollectionDocuments/Rapporten/Big_data_in_de_gezondheidszorg.pdf

Ottes, L. (2016). *Big data in de zorg.* The Hague: Wetenschappelijke raad voor het regeringsbeleid.

Over de UMC's. (2017). Retrieved from NFU: www.nfu.nl/umc

Paulk, M. (1993). Capability Maturity Model for Software. *Encyclopedia of Software Engineering*.

R, C., Knight, M., & Montgomery, A. (2012). *Maturity models 101 A primer for applying maturity models to smart grid security, resilience, and interoperability.* Pittsburgh: Carnegie Mellon University.

Raghupathi, W., & V, R. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems, 2.1*, 3.

Russom, P. (2011, Fourth Quarter). Big data analytics. *TDWI Best practices report*, pp. 1-35.

Salah, D., Paige, R., & Carins, P. (2014). An evaluation template for expert review of maturity models. *International Conference on Product-Focused Software Process Improvement*, 318-321.

Saltz, J., & Shamshurin, I. (2016). Big data team process methodologies: a literature review and the identification of key factors for a project's success. *IEEE International Conference* (pp. 2872-2879). IEEE.

Selanikio, J. (2016). *Big data and health: what is big data, how did it happen, and what does it mean for health?* Magpi. Retrieved from http://home.magpi.com/how-big-data-improved-ebola-control

Shaffer, V., & Craft, L. (2016). *Hype cycle for healthcare providers.* Retrieved from Gartner: http://www.garnter.com

Shih, W., & Chai, S. (2016, January). Data-Driven vs. Hypothesis-Driven Research: Making sense of big data. *Academy of Management Proceedings*.

StClaire, T. (2017, May 8). *How Blockchain can solve real problems in healthcare*. Retrieved from LinkedIn: https://www.linkedin.com/pulse/how-blockchain-can-solve-real-problems-healthcare-tamara-stclaire

Sun, J., & Reddy, C. (2013). Big data analytics in healthcare. *19th ACM SIGKDD International conference on Knowledge discovery and data mining* (pp. 1525-1525). ACM.

Sutherland, F. (n.d.). *Creative thinking and data analysis: who in their right mind?* Canberra: Australian National University.

Suzman, R., & Beard, J. (2011). *Global Health and Aging.* World Health Organization. Retrieved from http://www.who.int/ageing/publications/global_health.pdf

TDWI Big data maturity model and assessment tool. (n.d.). Retrieved from TDWI: https://tdwi.org/pages/maturity-model/big-data-maturity-model-assessment-tool.aspx

Ward, J., & Barker, A. (2013). *Undefined by data: A survey of big data definitions.* arXiv.org. Retrieved from https://arxiv.org/abs/1309.5821

Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and software technology*, 1317-1339.

What is Apache Hadoop? (n.d.). Retrieved from Apache Hadoop: http://hadoop.apache.org/

Wikipedia. (2017, 04 25). *Likert Scale*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Likert_scale

Williams, P., & Andersen. (2001). Information Security Governance. *Information Security Technical Report*, 60-70.

APPENDIX A: PRACTICAL EXAMPLES OF BIG DATA ANALYTICS IN HEALTHCARE

This appendix contains found practical examples of big data analytics in healthcare with some additional labels. For each of the examples a short background of the example is provided. Then, each of these examples is assessed on if they are really big data examples. Finally, they are labeled on what type of benefit it brings to the healthcare industry, as categorized by McKinsey (Groves, Kayyali, Knott, & Van Kuiken, 2013).

> **Structure of each example**
>
> **# Name of the example (source)**
> Short description of the example
> **Big data:**
> Does the example adhere to the adopted definition of big data? Why not?
> **Benefit:**
> In which of the five areas can this example be categorized: right living, right care, right value, right provider or right innovation

1 Heritage health prize (Heritage Health Prize, 2017)
Heritage Health had set the challenge to identify patients who will be admitted to a hospital within the next year using historical claims data. The goal was to identify patients at high-risk and ensure they get the treatment they need.
Big data: No, low variety of data
Benefit: Right living

2 Project Artemis (Cottle, et al., 2013)
Project Artemis uses real-time data from premature babies such as data from devices to predict infections 24 hours before symptoms appear.
Big data: Yes
Benefit: Right care

3 Joint-replacement (Cottle, et al., 2013)
Surgeons at Brigham and Women's hospital in Boston combine their own experience with data from research to systematically standardize knee-joint-replacement surgery with a resultant increase in more successful outcomes and reduced costs.
Big data: No, no high variety and no high velocity
Benefit: Right care, right value

4 Google Flu Trends (Cottle, et al., 2013)
Researchers at the Johns Hopkins School of Medicine found that they could use data from Google Flu Trends (a free, publicly available aggregator of relevant search terms) to predict surges in flu-related emergency room visits a week before warnings came from the Centers for Disease Control and Prevention.
Big data: Yes
Benefit: Right living

5 Propeller health (Ottenheijm, 2015)
Propeller Health, an American company, focusses on disease management for COPD patients. Patients place a sensor on their inhaler. This inhales registers when, where and how often the inhaler is used. Combining this data with forty other sources such as weather data, traffic and the quality of the air, they are now able to predict which patients are at risk.
Big data: Yes
Benefit: Right living, right value

6 Aurora health care (Ottenheijm, 2015)
Aurora Health care developed Smart Chart: a data warehouse that houses their data. They use this data to look at effects of length of stay, complications and readmissions to get an idea of the quality of care and lower the costs. Since the introduction of this system, using data to provide information, the number of readmissions has decreased with ten percent, saving the company millions of dollars.
Big data: Yes
Benefit: Right care, Right value

7 Center for Personalized Cancer Treatment (CPCT, 2017)
The CPCT analyzes a patients cancer cells and creates a profile out of this. They combine this data with data from previous cases and try to predict which treatment has the best outcome for this patient. This way, care will be personalized and patients will no longer have to deal with inefficient treatments.
Big data: No, low velocity, though using unstructured data.
Benefit: Right care

8 Watson for cancer treatments (Selanikio, 2016)
Watson can generate in about 15 minutes an analysis that would typically take months to come up with. It bases its analysis on a large database with many different data sources.
Big data: Yes
Benefit: Right care, right value

9 Fraud prevention system (Ottenheijm, 2015)
Centers for Medicare and Medicaid Services (CMS) created a Fraud Prevention System that looks at declaration patterns made known frauds. By coupling this with all declarations, data fraud worth of millions of dollars is detected each year, a significantly higher amount than before this system was in place.
Big data: No, as there is no high variety and this is structured data
Benefit: Right value

10 North York General Hospital (Raghupathi & V, 2014)
North York General Hospital uses real-time analytics to improve patient outcomes and gain greater insight into the operations of healthcare delivery. They reported to have implemented a scalable, real-time analytics application to provide multiple perspectives, including clinical, administrative, and financial data.
Big data: Yes
Benefit: Right care, right value

11 Care protocols (Raghupathi & V, 2014)
A large, unnamed, healthcare provider is analyzing data in the electronic medical record system with the goal of reducing costs and improving patient care. This includes unstructured data from physician notes, pathology reports and other sources. Big data analytics is used to develop care

protocols and case pathways and to assist caregivers in performing customized queries.
Big data: Yes
Benefit: Right care, Right value

12 Predicting complications for brain injuries (Raghupathi & V, 2014)
Columbia University Medical Center analyzed complex correlations of streams of physiological data related to patients with brain injuries. Their goal was to provide medical professionals with critical and timely information to aggressively treat complications. The advanced analytics is reported to diagnose serious complications as much as 48 hours sooner than previously in patients who have suffered a bleeding stroke from a ruptured brain aneurysm.
Big data: Yes
Benefit: Right care

13 Adverse drug effects (Raghupathi & V, 2014)
California-based Kaiser Permanente associated clinical data with cost data to generate a key data set, the analytics of which led to the discovery of adverse drug effects and resulted in the withdrawal of Vioxx from the market.
Big data: No, low velocity and low variety with only two types of data, though clinical data is semi-structured
Benefit: Right care, right value

14 Improving Ebola control (Selanikio, 2016)
One of the great obstacles in dealing with the Ebola outbreak was in monitoring the number and location of people infected with the virus. The US Centers for Disease Control and Prevention worked to build a more robust and widespread mobile surveillance network, using basic text messaging data collection instead of paper forms.
Big data: No
Benefit: Right care, right value

15 Parkinson's Disease App (Selanikio, 2016)
mPower, a mobile Parkinson disease study uses a mix of surveys and tasks that activate phone sensors to collect and track health and symptoms of Parkinson Disease progression – like dexterity, balance or gait. The goal was learn about variations of PD, to improve the way they describe and manage these variations and to learn if mobile devices can improve the quality of life for people with PD.
Big data: No
Benefit: Right care, right innovation

APPENDIX B: THE MATURITY MODEL

This appendix contains version 1.0, version 2.0 and version 2.1 of the maturity model.

MATURITY MODEL VERSION 1.0

This section consists of the documentation on the maturity model version 1.0.

MATURITY LEVELS

This model consists of five maturity levels, based on the CMMI. These five levels are initial, repeatable, defined, managed and optimized. The main characteristics of these levels are described:

Initial

This is the starting point for big data analytics. Big data analytics are not performed within the academic hospital. There is not strategy on big data analytics and there are no systems or protocols in place to support this.

Repeatable

Big data analytics is not yet something that is embedded within the organization or standard, but there are small ad-hoc projects that can serve as a proof of concept within the hospital. Big data analytics could be repeated, but there is no standard established. It can still be a time consuming process to gather data.

Defined

Big data analytics can be performed within the hospital and there is a standard business process. There are processes in place to facilitate big data analytics throughout the hospital, but these still require manual labor and can thus be time consuming.

Managed

Big data analytics are really embedded in processes of the academic hospital. Next steps are taken to prepare processes, protocols and IT to enable big data analytics on a national level.

Optimized

Big data analytics have reached the highest level of maturity for academic hospital in the Netherlands. It is possible to perform big data analytics on data generated by different healthcare providers. The data is findable and accessible, without having to go through manual processes that might be time-consuming. There is one process that is adopted by all academic hospitals and only one IT system supporting this. Importantly, big data analytics can be performed proactive and not only retrospective. This means that analysis are not only focused on proving a suspected theory or relation, but rather to discover new insights in the data.

DOMAINS AND ATTRIBUTES

The model consists of seven domains that altogether contribute to big data analytics in academic hospitals. These domains are strategy, governance, technology, data, organization, privacy and innovation. Each of these domains consists of domain specific attributes that contribute to making a big domain comprehensible and measurable.

1 Strategy

The domain strategy involves the presence of big data analytics awareness in the top management, and the extent to which big data analytics is considered in the hospital's strategy. This also involves processes, the extent to which big data analytics is exploited in the operational and decision making processes to achieve the hospital's strategy. Funding of big data projects is also considered.

2 Governance
The governance domain evaluates the extent to which organizational structures are in place to define roles and responsibilities, authority, control of big data analytics. This domain is defined by the presence of a data governance program, the ownership of data, the presence of a data architecture that makes data findable and the existence of explicit data definitions.

3 Information technology
The information technology domain is part of the core of big data management. This focusses on the technology that is required to extract data and generate knowledge from all data sources within a hospital. The technology domain focusses on the way data is stored, technical access to data, the applications that can be used on the data and the ETL process for all data that enters the data warehouse.

4 Data
Data is a domain that is also, besides technology, part of the core of big data analytics within the hospital. This domain focusses on the data that is generated by the hospital which can be used for big data analytics purposes. The attributes of this domain cover the extent to which data is used within the hospital for analytics, different data types that can be used for analytics, the quality of the data and the data lifecycle for datasets that are used for big data analytics.

5 Organization
The organization domain is defined by the presence of people with big data analytics skills, data stewardship throughout the hospital, the culture within the hospital on data-driven decision making and the attitude towards big data analytics. We also consider knowledge sharing on big data analytics throughout the academic hospital.

6 Privacy
The privacy domain is defined by four attributes. The attributes of this domain are privacy policy monitoring, privacy governance, privacy awareness within the hospital and technology to anonymize or pseudonymize data.

7 Innovation
The last domain of the model is the innovation domain. As mentioned earlier, big data analytics can be useful to discover new relations or theories within the data that can be found by exploring the data in a different way than the data is now mostly used: starting with a hypothesis and proving it with the data. This domain tries to capture the innovative culture of the hospital with four attributes. These are the ways that big data analytics is used: retrospective or proactive, the existence of innovation within the hospital and its role in national innovation practices. The composition of big data analytics teams is considered: are they multidisciplinary? Finally the stimulation of creativity for employees working on innovation is considered.

Scoring scheme
When all questions are answered, the maturity of the hospital can be determined.
Each of the questions has five answers. For every domain, add the numbers corresponding to the answers and then divide by the number of questions asked. This is rounded to once decimal to determine the current maturity of the hospital. Do the same for the answers on the future of the hospital. If you use the scoring form the scores will be calculated automatically.

This is repeated for all domains, and results in 7 maturity scores for the hospital on each domain in the current situation and 6 maturity scores for the hospital in the near future. There is not one single maturity level defined, because this takes away too much detail of the current situation.

Finally, the seven maturity scores are added together and divided by 7 (rounded to 1 decimal) to determine the overall maturity of the academic hospital.

The same is done for the future maturity score of the academic hospital.

Structure of the maturity model

When you finish the scoring form, there are three visualizations of the maturity. First we provide the overview of the big data maturity (Figure 1) and secondly the detailed maturity of the hospital now (Figure 2) and the detailed maturity of the hospital in two years' time.

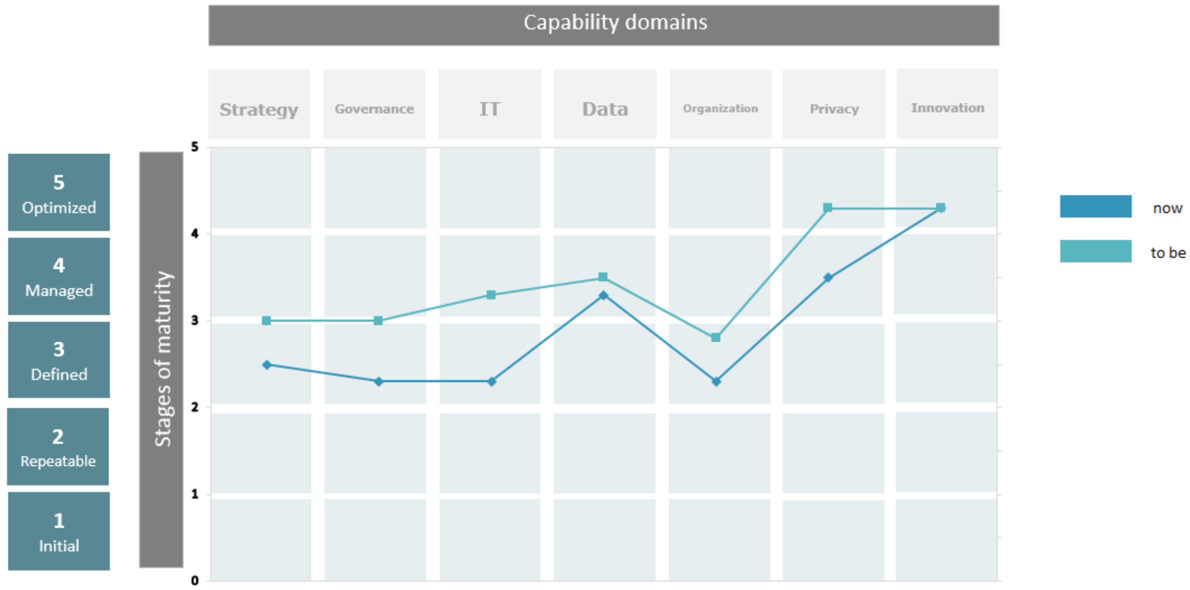


Figure 30 Big data maturity overview - example

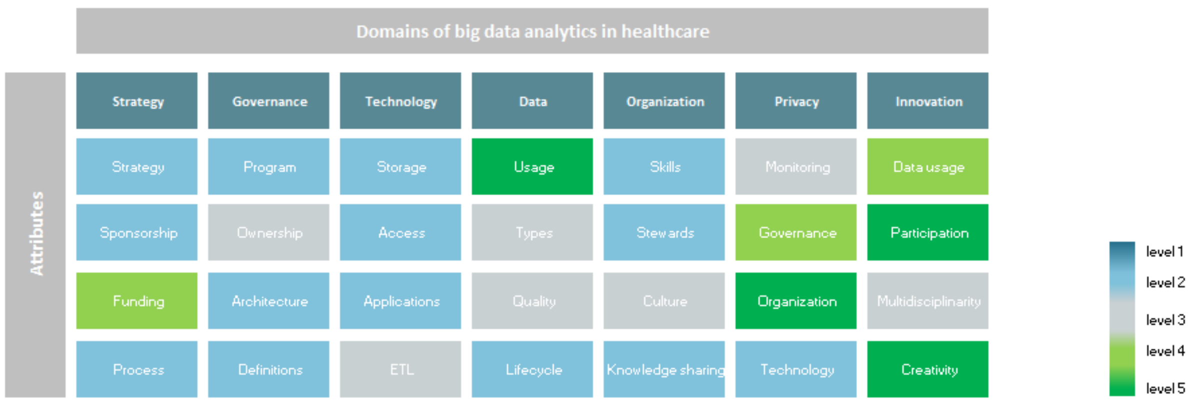


Figure 31 Big data maturity on detailed level now – example

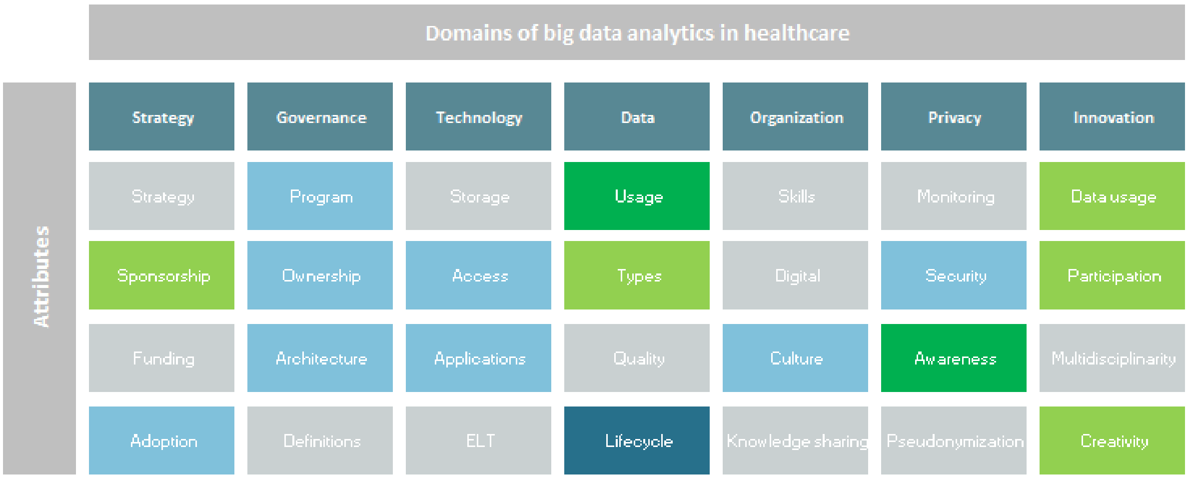


FIGURE 32 BIG DATA MATURITY ON DETAILED LEVEL TO BE - EXAMPLE

INSTRUCTIONS

Instructions on how to use the maturity model will be discussed in this section. Hospitals that want to assess their big data analytics maturity are advised to go through all steps described in this section.

Step 1: Role assignment

The hospital that wants to prepare for the assessment should first assign certain roles before it can start. One person should execute the assessment and define the employees that need to be involved, and one person high in the organization should sponsor the assessment.

Assessor

The assessor should execute the maturity assessment. The assessor needs to be familiar with big data and have enough experience within the hospital to know which persons to ask for the assessment. The assessor needs to collect data from IT and business. The assessor could be someone working on innovation or involved with research ict.

Sponsor

Since multiple people within the hospital will be asked to spend time on the assessment and the assessment might lead to change, someone high in the organization should sponsor the project. He or she should let the organization know that the project is important and that people should cooperate.

Step 2: Collecting the data

The next step is to start collecting the relevant data to answer the questions of the model. Roles that should be asked for input are (if present within the hospital):

Roles
IT architect
Head of Business Intelligence
Innovation manager
CIO
Privacy/ security officer
Researcher using big data analytics, or more advanced analytics in the hospital

The assessor should have enough knowledge of the hospital to know if these roles are present within the organization, and if not find suitable replacements.
Data can be collected by handing out the questions to the specific people chosen, or if there is not enough knowledge of big data analytics to let the questions directly be answered, interviews can also be sufficient.

Step 3: Determining the maturity
When there is enough data collected to answer all questions, the questions should be answered using the scoring form. When in doubt, the assessor should decide if more information is needed, or if he/she can make a decision himself. After finishing answering all the questions, the visualizations of the current maturity should be made.

A guideline to help answering the questions can be found in the next section.

Step 4: Interpretation
The model is a tool to quickly assess and illustrate the current maturity in big data analytics within the hospital. This is purely descriptive, and does not provide immediate improvement potentials. Therefore, interpretation of the outcomes is needed. There are a couple of steps to interpret the results:

1. As-is landscape
   The assessor should start by analyzing the current landscape within the organization. Define what the current limitations are, and how big data analytics is currently placed within the organization.
2. To-be landscape
   A hospital that wants to assess their current big data analytics maturity, most likely has goals concerning this topic. Make the to-be maturity clear for the academic hospital. For the to-be maturity a period of two years is advised.
3. Gap analysis
   Once both the as-is and the to-be landscape are clear, the assessor should analyze the gap between these landscapes. These gaps should be connected to domains of the maturity model.
4. Recommendations
   The gaps, and the related domains, should be translated by the assessor to actionable recommendations. This does not mean that the hospital should immediately strive for the highest maturity level, but to a reachable higher maturity that is aligned with the strategy of the hospital.

Step 5: Validation
To ensure that the assessor collected the right information and interpreted the information in the right way, the assessor should make a report and/or presentation to all persons that were involved. This should lead to a feedback session where the results might be slightly adjusted if needed. If a feedback session is difficult to organize, a questionnaire could be sufficient.

After the validation, the report and/or presentation should be finalized.

Step 6: Distribution of report

The final step of the maturity assessment is distribution of a report with the findings of the maturity assessment. Figure 4 shows a suggested table of contents for the report. After finishing the report, the assessor should actively spread the results within the hospital.

1. Table of contents
2. Introduction
3. Explanation of the model
    a. Maturity levels
    b. Domains and attributes
4. Research method
5. Results
    a. Strategic alignment
    b. Governance
    c. …
6. Conclusions and recommendations

FIGURE 33 EXAMPLE TABLE OF CONTENTS FOR THE REPORT

ASSESSMENT GUIDELINES AND QUESTIONS OF THE MODEL

Each of the 28 attributes is defined by a question with 5 answers: each corresponding with one maturity level. The questions of the model are discussed in this section, together with assessment guidelines for these questions. The attached scoring form is recommended to use as this automatically calculates the maturity and the graphical visualizations of the model.

Each question has five answers that progress stepwise in terms of maturity. The first answer corresponds to the first level of maturity and the fifth answer with the highest level of maturity. When answering the questions start with the first answer. If this answer fits the current situation best, stop there. Otherwise progress to the next answer. Continue this progress until the best fitting answer is found. When no answer matches the current situation, choose the best fitting answer.

Note that if the interviewees do not know the answer to the question or cannot offer insights on the domain or attribute, you can ask them for the people within the academic hospital that might be able to help you.

Strategic alignment

The questions for strategic alignment are typically answered by the CIO or comparable. Documents that can support you are the current documented strategy of the hospital and the organogram of the hospital.

Strategy	
1	Big data analytics are not considered for the strategy of the academic hospital
2	There is awareness on the possibilities of big data analytics but this is not documented in the strategy
3	There is a hospital wide big data strategy that is documented and accessible for every employee
4	There is a documented shared strategy on big data analytics between UMCs, accessible for everyone

5	There is a documented shared strategy on big data analytics between all researchers of medical data on a national level

Error! Not a valid link. Is there **sponsorship** for big data analytics in the highest management?
A sponsor promotes big data analytics and grants the mandate for the program. The sponsor continuously stretches the programs importance. We consider someone as a sponsor of big data analytics when this is embedded in the description of their role. Note that this is not always explicit.

Sponsorship	
A sponsor promotes big data analytics and grants the mandate for the program. The sponsor continuously stretches the programs importance.	
1	*No, there is no sponsorship for big data analytics within the academic hospital*
2	*There is no sponsorship for big data analytics within the management, but on a lower level in the academic hospital*
3	*The CIO (or someone comparable) is a sponsor for big data analytics*
4	*The CEO is a sponsor of big data analytics*
5	*Besides a sponsor within the board of the academic hospital, there is a sponsor for big data analytics in healthcare on a national level.*

Is there **funding** for big data programs?

Funding	
1	There is no funding for big data programs
2	There is funding for big data programs, but only for proof-of-concepts
3	There is funding for big data programs that are bigger than proof-of-concepts but comes mostly from the IT department
4	Big data programs are funded in executive and business unit levels.
5	Big data programs in the hospital are funded in executive and business unit levels and there is a shared budget between healthcare providers for big data purposes

Are big data analytics used in **processes**?
When big data analytics are actually used within the hospital, the repeatability of this process should be regarded.

This question could also be answered by someone responsible for research ICT or research ICT architecture, which can typically be found in an academic hospital in the Netherlands.

Processes	
When big data analytics are actually used within the hospital, the repeatability of this process should be regarded.	
1	Big data analytics is not used in processes
2	There are some siloed big data analytics efforts in individual departments
3	In most processes, big data analytics are used. Best practices are shared.
4	Big data analytics are used throughout the hospital and all following the same documented protocols.
5	Big data analytics are used by all healthcare providers following the same structured, documented protocols.

Governance

The questions for governance can mostly be answered by the head of the business intelligence department and the CIO.

Is there a data governance program in the hospital? When this is the case, you can read the documentation to formulate an answer to this question.

Data governance program	
"Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods."	
1	There is no data governance program
2	There is a data governance progam, but not formally
3	There is a formal data governance program within the hospital
4	There is a formal data governance program within the hospital and there is a national data governance program for shared healthcare data
5	There is one formal data governance program for all careproviders within the Netherlands

Data ownership is sometimes hard to find in big organizations such as academic hospitals. If data ownership is not explicit you can ask the head of the business intelligence who they contact for each of the data sources and if these contacts are aware of their responsibilities.

Data ownership	
Even though applications are often owned by IT, data ownership is often undefined. The owner of a business process is often also the owner of a related data asset. An owner has the highest level of responsibility over a specific data asset.	
1	Data ownership is not defined
2	Data ownership lies with IT
3	Data ownership lies with the business: each department is owner of their data
4	Data owners understand that they are accountable for this data and ownership is known or findable by other employees
5	Data ownership is clear for all data within the hospital and for data that is shared to other healthcare providers. At any point in the data chain it is known who the owner of data is. This is documented and accessible.

A data architecture can, when it exists within the academic hospital, most likely be found at the IT department or the business intelligence. If there is a data architecture, or something comparable such as a data catalogue then ask permission to view.

There is a **data architecture** available	
1	There is no data architecture available
2	There is a data architecture available, but only for some data sources within the hospital
3	There is a data architecture available, containing all data sources in the hospitals but only on a high level.
4	There is a data architecture available, containing all data sources in the hospital and extensive.
5	All data is FAIR: Findable, accessible, interoperable and readable from all academic hospitals to all academic hospitals. This is true for all (anonymized and not anonymized) medical data.

The question on data definitions can be best askes to the head of the business intelligence department. If it takes a lot of effort and time to add new data sources to the warehouse this might indicate a lack of data definitions. Low quality of data is also a good indicator.

Data **definitions**	
A data dictionary is a centralized repository of information about data such as meaning, relationships to other data, origin, usage and format.	
1	There are no data definitions
2	There are data definitions but they are not formalized
3	There are documented data definitions that are used throughout the UMC, but are not part of the standard proces
4	There is a data standard/protocol that is documented and is used for all data within the hospital that is entered in the central warehouse or data lake.
5	There is one data standard/protocol that is documented and adopted by all the UMCs in the Netherlands

Information technology

The questions for information technology can mostly be answered by the head of the business intelligence department.

Data storage	
1	All data is stored in siloes
2	There is a central data warehouse where most data is placed
3	All data is automatically integrated in the data warehouse. All data is accessible through the data warehouse.
4	There is one infrastructure that provides access to all stored data for UMCs
5	There is one infrastructure that provides access to all stored data for UMCs and is linked to all other healthcare providers i.e. biobank

To acquire the answers to the question on data access you could ask what the process would look like if you would be a scientist looking for data from the warehouse.

Data access to the data warehouse	
1	Data access is not always possible
2	Data access is possible, but it is a manual process
3	Data access is for most data sources provided through an automatic process for data within the hospital
4	Data access is an automatic process for all data sources within the hospital
5	There is an automatic process to get access to data within the hospital and data that is shared between hospitals

Continue on the previous subject, but now ask how you would analyze your data. Are there shared tools available?

Big data analytics **applications**	
1	There are no applications to perform big data analytics
2	There is some analytics software but this is siloed in a department

3	There is one hospital-shared interface with big data analytics applications, analytics performed by IT
4	Analytics can be performed by the whole company
5	There is a care provider-shared interface to all available analytics applications

The question on the ETL process tries to capture how data, when first entered in the data warehouse, is updated in the future. Is this still a manual process? For example data cleaning could still be a manual task to be executed.

ETL process: from database to data warehouse	
1	Data cannot be extracted from databases for analytics
2	There is an ETL process for all data, but this is done manually
3	ETL is done automatically for all data in the warehouse
4	ETL is done automatically for all data in the warehouse and there is a manual process to be able to retrieve medical data from other academic hospitals
5	ETL is done automatically for all data in the warehouse and there is an automatic process to be able to retrieve medical data from other academic hospitals

Data

The questions on data are most likely to be answered by the head of the business intelligence department. To get a proper view on the ways that data is now used within the hospital the person responsible for research ICT could either help answering it or get you in to contact with the researcher that is ahead of the rest in terms of big data analysis.

Usage	
1	Data is not used for analytics purposes
2	Data is used for reporting (dashboarding)
3	Data is used for monitoring (actual data)
4	Data is used for evaluation (why did it happen?)
5	Data is used for prediction

The question on data types will most likely be answered by the head of the BI department. Indicators might be: is free text accessible through the warehouse? How old is the data in the warehouse? Will patient-generated devices such as health apps be accessible in the warehouse?

Data **types** in the warehouse	
1	No data can be used for big data analytics
2	Some data can be used for big data analytics
3	Structured and unstructured data can be used for big data analytics
4	(Near) Real-time data is available
5	Patient-generated data is (near) real-time available for big data analytics

Most data warehouses have quality control somewhere in the process. Ask how data quality is improved.

Data **quality** of data in the warehouse

1	Data quality is not measured or maintained
2	Data quality is improved when the data is used for analytics purposes
3	All data is manually checked on quality before entering the central data warehouse
4	Data quality is updated automatically. Data stewards are responsible for data quality, not IT
5	Data quality is maintained with one standard for all data that is shared with care providers. Users trust that data quality is high enough for analytics without extensive quality improvement necessary.

When researchers or users of the data from the data warehouse have their data, it should still be maintained and eventually deleted from their workspace.

Data **lifecycle for big data sets**	
Just like any product, data typically goes through a number of stages. It is created, used, needs maintenance, back-ups, and eventually deletion or archiving.	
1	Lifecycle aspects are not considered for this dataset
2	Lifecycle aspects are considered on ad-hoc basis.
3	Data is maintained and archived manually on frequent basis for all datasets in the academic hospital
4	Data is maintained and archived automatically for all datasets in the hospital
5	Data is maintained and archived automatically for all shared data as well

Organization

The questions on organization are most likely to be answered by either the CIO, the head of the business intelligence or the person responsible for research ICT. It should be someone within the academic hospital with an overview of the users of data analytics within the hospital. Indicators for the question on skills could be: does every department have the role of data scientist or big data analyst fulfilled? Who is doing advanced data analytics?

Skills	
1	We have a lack of staff that have big data capabilities
2	We have some people that have big data analytics skills, but this is not their main role
3	There are data scientists that are the only ones using big data analytics
4	Within each department, there are business analysts that understand and use big data analytics
5	Staff feels empowered to experiment with big data tools beyond the formal definitions of their role. Experiences are shared within the hospital.

An indicator to answer the question on data stewardship would be to ask who would update the data quality of the original data source when this is low.

Data **stewardship**	
A data steward is a person responsible for the management of data quality for a certain data source.	
1	Data is considered as part of IT
2	Data is still part of IT, but data stewardship is recognized as important and data stewards are sought within departments

3	There are data stewards within each business unit that are responsible for their data management
4	There is a data board that harmonizes big data efforts throughout the hospital together with data stewards. There is a chief data officer in place that is accountable for all data. Data is no longer considered as purely IT.
5	There is a central steering committee that work with data stewards and the board to harmonize big data efforts throughout the hospitals

The question on culture could be difficult to answer. The business intelligence will most likely know which departments are using the data in their decision making processes.

Culture: Is data-driven decision making adopted in processes within the hospital?	
1	There is resistance towards big data initiatives or data driven decision making
2	The IT department uses data driven decision making in processes
3	The board uses data driven decision making in processes
4	Within the entire hospital data-driven decision is accepted and used. It is adopted in processes.
5	Data driven decision making is accepted and used. It is adopted in processes on a national level.

The question on knowledge sharing could be answered by asking the researcher using the most advanced data analytics techniques how their department shares their knowledge with other departments.

Knowledge sharing	
1	There is no knowledge on big data analytics within the hospital
2	Some individuals within the hospital have knowledge of big data analytics
3	All data scientist or employees with knowledge of big data analytics share their knowledge consequently
4	Departments or teams that want to learn about big data analytics can get educated by more advanced teams within the hospitals in an organized way
5	All care providers are educated by a central education program and there is a national platform to share knowledge and best practices

Privacy

The questions on privacy are most likely to be answered by the privacy or security officer and the head of the business intelligence. The CIO could also be your contact for these questions. Specifically ask how security on privacy is guaranteed for data in the data warehouse.

Is **monitoring** of compliance with privacy policy part of the organization for the data warehouse?	
1	Monitoring of compliance with privacy policies is not part of the organization
2	Compliance of privacy policies is monitored ad-hoc
3	Compliance of privacy policies is monitored manually, but consistently
4	Privacy policies are monitored automatically and flagged when necessary
5	There is a privacy board that is responsible for monitoring privacy policies on data between parties

Governance	
1	Nobody is accountable for security of data
2	Nobody is formally accountable for security of data but tasks are being performed
3	Responsibilities on data security are formally defined, but lie within IT
4	IT security is a joint responsibility of business and IT, and integrated with hospital strategy
5	There is external governance on the security and privacy of data for all data between hospitals

Indicators for organization could be to ask the privacy officer how awareness is created on privacy issues.

Organization	
1	There is little awareness of privacy issues
2	Privacy of patient data is discussed when employee is hired
3	There are documents available for education purposes on privacy of patient data
4	Within each department there is someone responsible for training and maintaining knowledge on privacy of the department
5	Privacy is on top of mind for all employees that are dealing with data

For the question on technology ask the head of business intelligence how and when in the process anonymization or pseudonymisation is done.

Technology	
1	There is no technical solution for anonymization or pseudonymisation of data
2	There is a technical solution for anonymization or pseudonymisation of data but has to be executed through a manual process
3	Anonymization or pseudonymisation of all data can be done through an automatic process
4	Privacy by design is the used approach throughout the whole engineering process
5	Data from different care providers can be shared, while for anonymization or pseudonymisation of this data is in place

Innovation

The questions on innovation are most likely to be answered by someone responsible for innovation within the academic hospital. The CIO and the person responsible for research ICT might offer answers to the questions.

Usage of big data	
You can use big data to prove theories or use to discover new theories or correlations. To prove a theory, you might not need big data after all.	
1	We do not use data analytics
2	We do use analytics, but nog on big data
3	We use big data analytics to prove suspeted theories (retro-spective)
4	We use big data analytics not only retrospective, but also proactive: the data gives us new insights and we can explore this data to come up with new theories
5	We use big data analytics not only retrospective, but also proactive and combine our data with other datasources such as weather data to discover unsuspected patterns

To know how actively the UMC participates in innovation on a national level, ask the CIO or innovation team what their role is in the national initiatives already existing such as the meetings set-up by the NFU (Nederlandse Federatie van Universitair Medische Centra)

The UMC actively **participates** in innovation	
1	Innovation is not part of the academic hospital
2	Innovation is part of certain departments of the academic hosptial
3	Innovation is hospital-wide part of the academic hospital
4	Besides innovation being a part of the entire UMC, the UMC participates in innovation programs on a national level
5	Besides innovation being a part of the entire UMC, the UMC has a leading role in innovation programs on a national level

Multidisciplinary teams	
1	Big data analytics are not performed
2	Big data analytics are performed purely by individuals
3	Big data analytics are performed purely by individuals, but do share experiences with others in their team
4	Big data analytics are performed purely by individuals or teams, but share their knowledge and best practices horizontally across the organization
5	Big data analytics are performed by multidisciplinary teams

Creativity: There is time to think creatively for the innovation team	
1	We do not have people that have a role involving innovation
2	There is no time for creativity for people involved with innovation in their day-to-day conduct of work
3	Some working hours could be spent on creative ideas, but this is not encouraged actively
4	Employees can spend time on creativiy and this is encouraged actively by their manager
5	Employees can spend time on creativity and this is encouraged actively by their manager. Employees are educated on creative thinking techniques or 'design thinking'

MATURIY MODEL VERSION 2.0

This section consists of the documentation on the maturity model version 2.0

BIG DATA

This model uses the following definition for big data and big data analytics:

Definition: Big data

Big data is defined as a high volume of data that can no longer be stored using traditional methods, a high variety of data combining both structured and unstructured data and a high velocity of data: the data set is constantly changing.

Definition: Big data analytics

Big data analytics is where analytical techniques are used to operate on big data with a data-driven approach for discovery analytics or exploratory analytics. The data is used to explore, not to prove existing hypotheses.

MATURITY LEVELS

This model consists of five maturity levels, based on the CMMI. These five levels are initial, repeatable, defined, managed and optimized. The main characteristics of these levels are described:

Initial
This is the starting point for big data analytics. Big data analytics are not performed within the academic hospital. There is not strategy on big data analytics and there are no systems or protocols in place to support this.

Repeatable
Big data analytics is not yet something that is embedded within the organization or standard, but there are small ad-hoc projects that can serve as a proof of concept within the hospital. Big data analytics could be repeated, but there is no standard established. It can still be a time consuming process to gather data.

Defined
Big data analytics can be performed within the hospital and there is a standard business process. There are processes in place to facilitate big data analytics throughout the hospital, but these still require manual labor and can thus be time consuming.

Managed
Big data analytics are really embedded in processes of the academic hospital. Next steps are taken to prepare processes, protocols and IT to enable big data analytics on a national level.

Optimized
Big data analytics have reached the highest level of maturity for academic hospitals in the Netherlands. It is possible to perform big data analytics on data generated by different healthcare providers. The data is findable and accessible, without having to go through manual processes that might be time-consuming. There is one process that is adopted by all academic hospitals and only one IT system supporting this. Importantly, big data analytics can be performed proactive and not only retrospective. This means that analysis are not only focused on proving a suspected theory or relation, but rather to discover new insights in the data.

Domains and attributes

The model consists of seven domains that altogether contribute to big data analytics in academic hospitals. These domains are strategy, governance, technology, data, organization, privacy and innovation. Each of these domains consists of domain specific attributes that contribute to making a big domain comprehensible and measurable.

1 Strategic alignment

Without the support of the full organization, big data analytics will not succeed. The board has to set out a strategy that will define *how* big data analytics will be used within the organization. This should be defined in a strategic document. A clear big data strategy is considered as key to successful adoption of big data analytics within an organization. This strategy has to be adopted by the whole organization to make this succeed. The strategy needs to formulate a clear vision, obtaining the buy-in within the whole organization and not only IT. Sponsorship also involves funding and an advocate of the program in the board of the hospital.

2 Governance

Big data governance formulates policy relating to optimization, privacy and monetization of big data by aligning the objectives of multiple functions. These policies are on metadata (setting definitions for data), access (who gets access to data?), data ownership, data quality, data security, data assets and data lifecycle. Governance does not consider data on operational level but sets guidelines and rules on how to use the data, and who is responsible for what in the organization considering big data analytics. Data ownership should be defined for every data source at each point in the big data analytics process.

3 Information technology

Big data analytics is in essence a technological solution. In the perfect world, all data sources should be connected to a central data warehouse. This includes all internal data sources, but also applications that are used by patients or medical monitoring devices such as heartrate monitors. To ensure the best care for the patient, all relevant patient data should be available. The central data storage should be able to deal with the volume, velocity and variety of big data.

4 Data

Data is a domain that is also, besides technology, part of the core of big data analytics within the hospital. To get the best result out of big data analytics, data should be of high quality. Data can be used for many different purposes. Five ways of using data are distinguished, from low to high maturity, reporting, monitoring, evaluation and finally prediction. All data should be available for data analysis. So not only structured data, but also unstructured data and (near) real-time data. Besides the internal data sources of the UMC, external data sources should be considered such as weather data.

5 Organization

A big data analytics project will only succeed when the right people are hired or trained to do this job. The organization should be as digital as possible. Big data analytics can only be performed on digital data. If most data is still in paper documents, this data cannot be used. The industry is still hesitant towards big data analytics as it is not the adopted way of doing research, which is hypothesis-driven. The adoption of data-driven decision making is considered in this domain. Finally, academic hospitals are almost a small town on their own. They all employ around seven to ten thousand employees. In an ideal world, these employees would share their knowledge on big data analytics so that the wheel is not reinvented again and again.

6 Privacy
Privacy issues have become increasingly urgent as more and more personal data is online. The electronic patient files are special because there are concerns of patients about disclosure of personal health information to third parties such as insurers or employers. However, the data should be shared and not kept within silos, so there has to be a mutual understanding between data sharing while keeping privacy and security standards high. Ideally, the data in a central data storage that is used for analytics will never leave a secure environment such as a so-called *sandbox.* Security audits should be in place and there should be frequent checks to see if the situation meets compliance and legislations standards. This also involves training users of this data on the matter. Data should be shared, but not without protection.

7 Innovation
The last domain of the model is the innovation domain. As mentioned earlier, big data analytics can be useful to discover new relations or theories within the data that can be found by exploring the data in a different way than the data is now mostly used: starting with a hypothesis and proving it with the data. Innovation is closely related to a climate of creativity and the composition of teams. Multidisciplinary teams have a higher capability of thinking outside the box.

Scoring scheme
When all questions are answered, the maturity of the hospital can be determined.
Each of the questions has five answers. For every domain, add the numbers corresponding to the answers and then divide by the number of questions asked. This is rounded to once decimal to determine the current maturity of the hospital. Do the same for the answers on the future of the hospital. If you use the scoring form the scores will be calculated automatically.

This is repeated for all domains, and results in 7 maturity scores for the hospital on each domain in the current situation and 7 maturity scores for the hospital in the near future. There is not one single maturity level defined, because this takes away too much detail of the current situation.

Finally, the seven maturity scores are added together and divided by 7 (rounded to 1 decimal) to determine the overall maturity of the academic hospital.

The same is done for the future maturity score of the academic hospital.

Structure of the maturity model
When you finish the scoring form, there are three visualizations of the maturity. First we provide the overview of the big data maturity (Figure 1) and secondly the detailed maturity of the hospital now (Figure 2) and the detailed maturity of the hospital in two years' time.

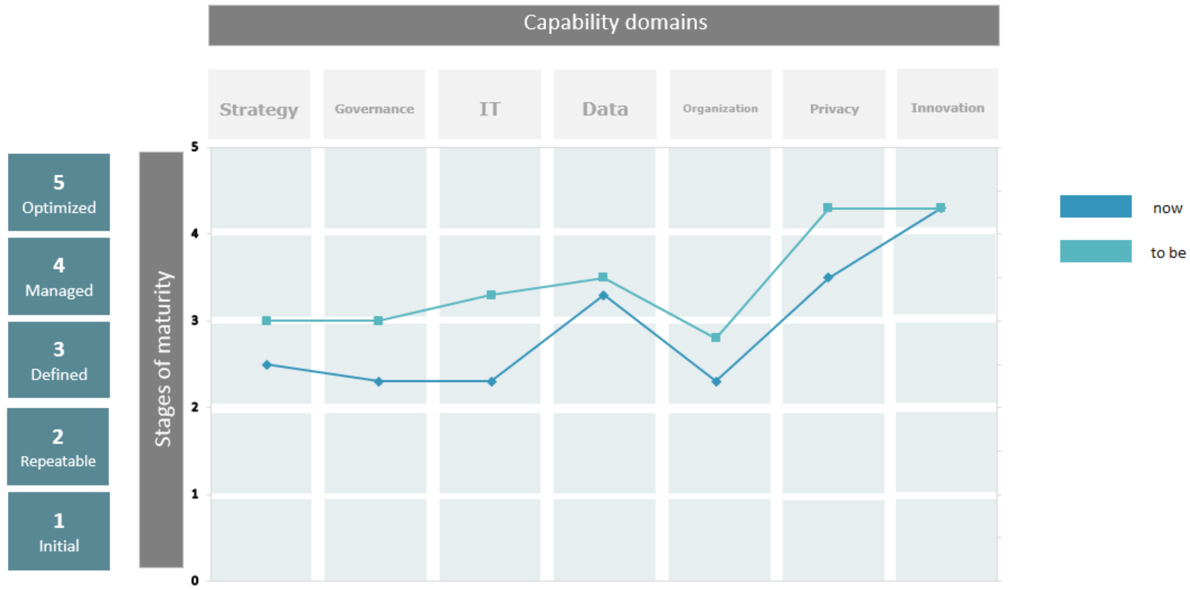


FIGURE 34 BIG DATA MATURITY OVERVIEW - EXAMPLE

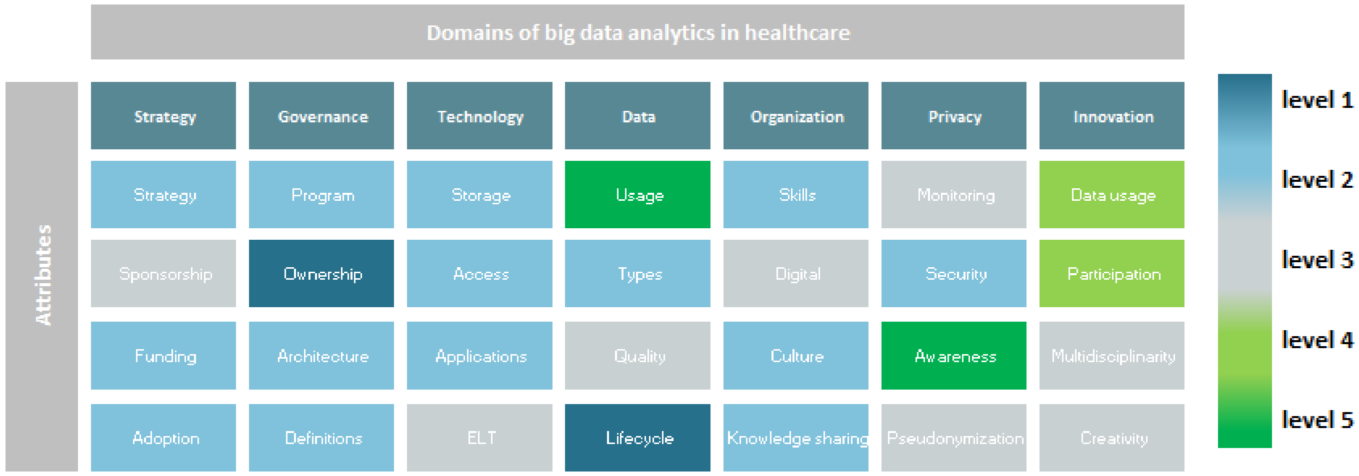


FIGURE 35 BIG DATA MATURITY ON DETAILED LEVEL NOW – EXAMPLE

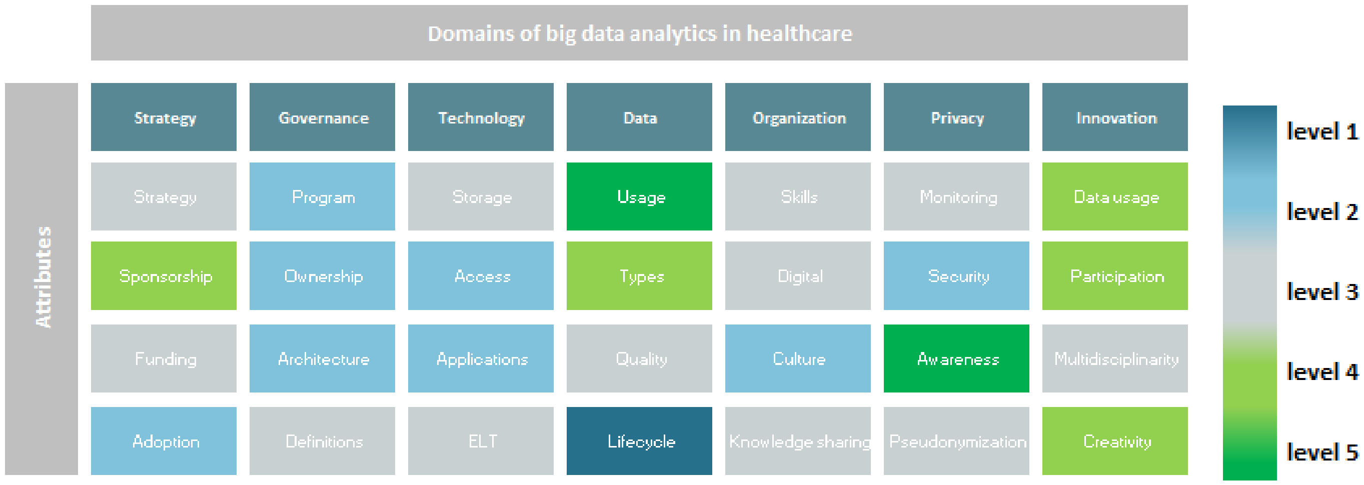


FIGURE 36 BIG DATA MATURITY ON DETAILED LEVEL TO BE - EXAMPLE

INSTRUCTIONS

Instructions on how to use the maturity model will be discussed in this section. Hospitals that want to assess their big data analytics maturity are advised to go through all steps described in this section.

Step 1: Role assignment

The hospital that wants to prepare for the assessment should first assign certain roles before it can start. One person should execute the assessment and define the employees that need to be involved, and one person high in the organization should sponsor the assessment.

Assessor

The assessor should execute the maturity assessment. The assessor needs to be familiar with big data and have enough experience within the hospital to know which persons to ask for the assessment. The assessor needs to collect data from IT and business. The assessor could be someone working on innovation or involved with research ict.

Sponsor

Since multiple people within the hospital will be asked to spend time on the assessment and the assessment might lead to change, someone high in the organization should sponsor the project. He or she should let the organization know that the project is important and that people should cooperate.

Step 2: Collecting the data

The next step is to start collecting the relevant data to answer the questions of the model. Roles that should be asked for input are (if present within the hospital):

Roles
IT architect
Head of Business Intelligence
Innovation manager
CIO
Privacy/ security officer
Researcher using big data analytics, or more advanced analytics in the hospital

The assessor should have enough knowledge of the hospital to know if these roles are present within the organization, and if not find suitable replacements.
Data can be collected by handing out the questions to the specific people chosen, or if there is not enough knowledge of big data analytics to let the questions directly be answered, interviews can also be sufficient.

Step 3: Determining the maturity

When there is enough data collected to answer all questions, the questions should be answered using the scoring form. When in doubt, the assessor should decide if more information is needed, or if he/she can make a decision himself. After finishing answering all the questions, the visualizations of the current maturity should be made.

A guideline to help answering the questions can be found in the next section.

Step 4: Interpretation

The model is a tool to quickly assess and illustrate the current maturity in big data analytics within the hospital. This is purely descriptive, and does not provide immediate improvement potentials.

Therefore, interpretation of the outcomes is needed. There are a couple of steps to interpret the results:

1. As-is landscape
   The assessor should start by analyzing the current landscape within the organization. Define what the current limitations are, and how big data analytics is currently placed within the organization.
2. To-be landscape
   A hospital that wants to assess their current big data analytics maturity, most likely has goals concerning this topic. Make the to-be maturity clear for the academic hospital. For the to-be maturity a period of two years is advised.
3. Gap analysis
   Once both the as-is and the to-be landscape are clear, the assessor should analyze the gap between these landscapes. These gaps should be connected to domains of the maturity model.
4. Recommendations
   The gaps, and the related domains, should be translated by the assessor to actionable recommendations. This does not mean that the hospital should immediately strive for the highest maturity level, but to a reachable higher maturity that is aligned with the strategy of the hospital.

Step 5: Validation
To ensure that the assessor collected the right information and interpreted the information in the right way, the assessor should make a report and/or presentation to all persons that were involved. This should lead to a feedback session where the results might be slightly adjusted if needed. If a feedback session is difficult to organize, a questionnaire could be sufficient.

After the validation, the report and/or presentation should be finalized.

Step 6: Distribution of report
The final step of the maturity assessment is distribution of a report with the findings of the maturity assessment. Figure 4 shows a suggested table of contents for the report. After finishing the report, the assessor should actively spread the results within the hospital.

1. Table of contents
2. Introduction
3. Explanation of the model
   a. Maturity levels
   b. Domains and attributes
4. Research method
5. Results
   a. Strategic alignment
   b. Governance
   c. …
6. Conclusions and recommendations

FIGURE 37 EXAMPLE TABLE OF CONTENTS FOR THE REPORT

ASSESSMENT GUIDELINES AND QUESTIONS OF THE MODEL

Each of the 28 attributes is defined by a question with 5 answer: each corresponding with one maturity level. The questions of the model are discussed in this section, together with assessment guidelines for these questions. The attached scoring form is recommended to use as this automatically calculates the maturity and the graphical visualizations of the model.

Each question has five answers that progress stepwise in terms of maturity. The first answer corresponds to the first level of maturity and the fifth answer with the highest level of maturity. When answering the questions start with the first answer. If this answer fits the current situation best, stop there. Otherwise progress to the next answer. Continue this progress until the best fitting answer is found. When no answer matches the current situation, choose the best fitting answer.

Note that if the interviewees do not know the answer to the question or cannot offer insights on the domain or attribute, you can ask them for the people within the academic hospital that might be able to help you.

Strategic alignment

The questions for strategic alignment are typically answered by a C-level executive. Documents that can support you are the current documented strategy of the hospital and the organogram of the hospital.

Strategy	
1	Big data analytics are not considered for the strategy of the academic hospital
2	There is awareness on the possibilities of big data analytics but this is not documented in the strategy
3	There is a hospital wide documented big data strategy
4	There is a hospital wide documented big data strategy that is actionable, i.e. with a roadmap
5	Besides a hospital wide strategy on big data, there is a documented shared strategy on big data analytics between UMCs that is actionable, i.e. with a roadmap

Is there **sponsorship** for big data analytics in the highest management?
A sponsor promotes big data analytics and grants the mandate for the program. The sponsor continuously stretches the programs importance. We consider someone as a sponsor of big data analytics when this is embedded in the description of their role. Note that this is not always explicit.

Sponsorship	
A sponsor promotes big data analytics and grants the mandate for the program. The sponsor continuously stretches the programs importance.	
1	*No, there is no sponsorship for big data analytics within the academic hospital*
2	*There is no sponsorship for big data analytics within the management, but on a lower level in the academic hospital*
3	C-level executive is a sponsor for big data analytics
4	*The CEO is a sponsor of big data analytics*
5	*Besides a sponsor within the board of the academic hospital, there is a sponsor for big data analytics in healthcare on a national level.*

Is there **funding** for big data programs?

Funding	
1	There is no funding for big data programs
2	There is funding for big data programs, but only for proof-of-concepts
3	There is funding for big data programs that are bigger than proof-of-concepts but this budget comes mostly from the IT department
4	Big data programs are funded in executive and business unit levels.
5	Big data programs in the hospital are funded in executive and business unit levels and there is a shared budget between healthcare providers for big data purposes

Adoption

The purpose of this question is to measure if big data analytics are used within the hospital. Also, when this is adopted the repeatability of the process is regarded: are there protocols and guidelines to structure these analysis and make them repeatable?
This question could also be answered by someone responsible for research ICT or research ICT architecture, which can typically be found in an academic hospital in the Netherlands.

Adoption	
When big data analytics are actually used within the hospital, the repeatability of this process should be regarded.	
1	Big data analytics is not used
2	There are some siloed big data analytics efforts in individual departments
3	In most departments, big data analytics are used. The projects do not follow one protocol
4	Big data analytics are used throughout the hospital, using the same protocol
5	Big data analytics are used by all UMCs following the same structured protocols.

Data Governance

The questions for governance can mostly be answered by the head of the business intelligence department and the CIO. Big data governance is measured because a mature big data governance program is a precondition for mature big data analytics.

Is there a big data governance program in the hospital? When this is the case, you can read the documentation to formulate an answer to this question.
The purpose of this question is to capture if there is a data governance program, and if this also captures big data.

Big data governance program	
"Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods."	
1	There is no big data governance program
2	There is awareness on a big data governance program
3	There is a big data governance program, but not formally documented
4	There is a formal big data governance program within the hospital
5	There is one formal big data governance program for all UMCs within the Netherlands

Data ownership is sometimes hard to find in big organizations such as academic hospitals. If data ownership is not explicit you can ask the head of the business intelligence who they contact for each of the data sources and if these contacts are aware of their responsibilities.

Data ownership	
Even though applications are often owned by IT, data ownership is often undefined. The owner of a business process is often also the owner of a related data asset. An owner has the highest level of responsibility over a specific data asset.	
1	Data ownership is not defined
2	Data ownership lies with IT
3	Data ownership lies with the business: each department is owner of their data
4	Data ownership lies with the business and owners are aware of responsibilities as owner
5	Data ownership is clear for all data in the hospital, and for all data shared between UMCs

A data architecture can, when it exists within the academic hospital, most likely be found at the IT department or the business intelligence. If there is a data architecture, or something comparable such as a data catalogue then ask permission to view. A data architecture should provide a clear overview of all data that is currently available for big data analysis. In the ideal situation not only data sources are known, but also metadata on this data. The data architecture should at least be available for the central data storage such as a data warehouse or a data lake.

There is a **data architecture** available for the **central data storage**	
1	There is no data architecture available
2	There is a data architecture available, but it is not complete
3	There is a complete data architecture available, but does not contain metadata
4	There is a complete data architecture available, including metadata
5	All data within the central data storage is FAIR: Findable, accessible, interoperable and readable

The question on data definitions can be best askes to the head of the business intelligence department. If it takes a lot of effort and time to add new data sources to the warehouse this might indicate a lack of data definitions. Goal is to have unambiguous data in the central data storage that is understandable for all users by reading the data dictionary.

Data **definitions** for the **central data storage**	
A data dictionary is a centralized repository of information about data such as meaning, relationships to other data, origin, usage and format.	
1	There are no data definitions used for data in the central data storage
2	There are data definitions but they are not formalized
3	There are documented data definitions that are used for data in the central data storage, for all data in the central data storage
4	There are documented data definitions that are used for data in the central data storage, for all data in the central data storage
5	There is one standard set of data definitions that are used by all UMCs in the Netherlands for data in the central data storage

Information technology

The questions for information technology can mostly be answered by the head of the business intelligence department.

The purpose of the question on data storage is that siloed data is often hard to extract or use for big data analytics. To make big data analytics available for everyone, data should be made available for everyone. One central data storage is a way to accomplish this.

Data storage	
1	All data is stored in siloes
2	Data is collected in central data storages
3	There is one central data storage where most data sources are accessible
4	There is one central data storage where all data sources are accessible
5	There is one central data storage where all data sources are accessible and one accepted infrastructure to access data from other UMCs

To acquire the answers to the question on data access you could ask what the process would look like if you would be a scientist looking for data from the central data storage.
The purpose of this question is to determine the ease with which an employee would have access to data. In the ideal situation this should not take time and no intervention from the department managing the central data storage should be necessary.

Data access to the **central data storage**	
1	Data access is not always possible
2	Data access is sometimes possible
3	Data access is possible, but is a manual process executed by the department responsible for the central data storage
4	Data access is possible, this process is automatic but you need the assistance of the department responsible for the central data storage
5	Automatic data access is possible without the need for assistance of the department responsible for the central data storage

Continue on the previous subject, but now ask how you would analyze your data. To take away concerns in terms of privacy or security of data, one secured environment where data is placed is something that is discussed by UMCs.

Big data analytics **applications**	
1	There are no applications to perform big data analytics
2	There are applications to perform big data analytics with, but this is individually bought by a department
3	There is a digital environment with applications where data could be analyzed
4	There is one digital environment with applications from the UMC where all data is analyzed
5	There is one digital environment with applications where data from multiple UMCs is analyzed

The question on the ETL process tries to capture how data, when first entered in the data warehouse, is updated in the future. Is this still a manual process? For example data cleaning could still be a manual task to be executed.

ETL process focus on extracting data from data sources, than transform the data to a certain set of rules or functions and then to load it into a central data storage which is often a data warehouse.

However, with the rise of big data, these methods often take too much time and effort and is high maintenance. The solution for this, comes with the rise of the Data Lake and ELT. Extract all data from the data sources, store it raw into a data lake and then transform the data dependent of the request.

The purpose of this question is that it should take the least time possible to get data from the source to the user.

from ETL to **ELT:** from data source to central data storage	
1	Data cannot be extracted from databases for analytics
2	There is an ETL process for all data, but this is done manually
3	ETL is done automatically for all data in the warehouse
4	ELT is done automatically for all data in the data lake of the UMC
5	ELT is done automatically for all data in the data lake of the UMC and there is an ELT standard to access other UMCs accessible data

Data

The questions on data are most likely to be answered by the head of the business intelligence department. To get a proper view on the ways that data is now used within the hospital the person responsible for research ICT could either help answering it or get you in to contact with the researcher that is ahead of the rest in terms of big data analysis.

Usage	
1	Data is not used for analytics purposes
2	Data is used for reporting (dashboarding)
3	Data is used for monitoring (actual data)
4	Data is used for evaluation (why did it happen?)
5	Data is used for prediction

The question on data types will most likely be answered by the head of the BI department. Indicators might be: is free text accessible through the warehouse? How old is the data in the warehouse? Will patient-generated devices such as health apps be accessible in the warehouse?

Data **types** in the central data storage	
1	No data can be used for big data analytics
2	Structured data is available
3	Structured and unstructured data is available, including free text
4	(Near) Real-time data is available
5	External data sources (data that is not generated by the hospital) are available

Most central data storages have quality control somewhere in the process. Ask how data quality is improved.

Data **quality** of data in the central data storage	
1	Data quality is not measured or maintained

2	Data quality is improved of the data extracted from central data storage
3	Data quality is improved of the data in the central data storage
4	Data quality is improved of the data in the original data source
5	There is one standard that defines data quality for all data shared between UMCs

When researchers or users of the data from the data warehouse have their data, it should still be maintained and eventually deleted from their workspace.

Data **lifecycle for big data sets**	
Just like any product, data typically goes through a number of stages. It is created, used, needs maintenance, back-ups, and eventually deletion or archiving.	
1	Lifecycle aspects are not considered for this dataset
2	Lifecycle aspects are considered on ad-hoc basis.
3	Lifecycle aspects are manually considered on frequent basis
4	Lifecycle aspects are automatically considered on frequent basis
5	Lifecycle aspects are automatically considered on frequent basis for all data in the UMC and shared with other UMCs

Organization

The questions on organization are most likely to be answered by either the CIO, the head of the business intelligence or the person responsible for research ICT. It should be someone within the academic hospital with an overview of the users of data analytics within the hospital. Indicators for the question on skills could be: does every department have the role of data scientist or big data analyst fulfilled? However, in the ideal situation these roles should not be fulfilled by employees within a department as this may narrow the scope of these project immensely.

Skills	
1	We have a lack of staff that have big data capabilities
2	We have some people that have big data analytics skills, but this is not their main role
3	In some departments there are people with big data analytics as part of their role
4	Within each department there are people with big data analytics as part of their role
5	There are people with big data analytics that can work on projects that are bigger than within the scope of one department

A precondition for big data analytics to exist in an academic hospital is to have digital data. Most academic hospitals have a HIMMS EMRAM[2] certification. A hospital that is a stage 7 EMRAM organization is paperless. This question could be answered by the CIO of the UMC.

How **digital** is the UMC?	
A completely digital UMC does not use any form of paper	
1	The UMC is not digital
2	The UMC has an electronic patient record system
3	The UMC handles most data electronically

[2] http://www.himss.eu/healthcare-providers/emram

4	The UMC is completely paperless
5	The UMC is completely paperless and there is no *shadow-IT in the UMC present*

The question on culture could be difficult to answer. The business intelligence will most likely know which departments are using the data in their decision making processes and can give an estimation (or a precise percentage) of the number of departments using big data analytics.

Culture: How many departments are performing big data analytics with data from the central data storage?	
1	There are no departments that use big data analytics with data from the central data storage
2	Some departments in the UMC are using big data analytics on data from the central data storage, but this is not tracked
3	The UMC tracks the number of departments that are performing big data analytics on data from the central data storage
4	Over 50% of all departments in the UMC are performing big data analytics on data from the central data storage
5	Over 50% of all departments in the UMC are performing big data analytics on data that is shared between UMCs

The question on knowledge sharing could be answered by asking the researcher using the most advanced data analytics techniques how their department shares their knowledge with other departments.

Knowledge sharing	
1	There is no knowledge on big data analytics within the hospital
2	Some individuals within the UMC have knowledge of big data analytics
3	Knowledge on big data analytics is infrequently shared in the UMC
4	Knowledge on big data analytics is frequently shared in the UMC
5	Knowledge on big data analytics is frequently shared in the UMC and also exchanged between UMCs

Privacy

The questions on privacy are most likely to be answered by the privacy or security officer and the head of the business intelligence. The CIO could also be your contact for these questions. Specifically ask how security on privacy is guaranteed for data in the data warehouse.

Is **monitoring** of compliance with privacy policy part of the organization for the data warehouse?	
1	Monitoring of compliance with privacy policies is not part of the organization
2	Compliance of privacy policies is monitored ad-hoc
3	Compliance of privacy policies is monitored manually, but consistently
4	Privacy policies are monitored automatically and flagged when necessary
5	There is a privacy board that is responsible for monitoring privacy policies on data between parties

Being able to experiment with big data and queries in a safe and secure "sandbox" test environment is important to both IT and end business users as companies get going with big data. Being able to experiment with big data and queries in a safe and secure "sandbox" test environment is important to both IT and end business users as companies get going with big data. If data is only accessible via a secured “sandbox”, data privacy and protection can be provided.

Security: is the data used for big data analytics from the central data storage used in a secure way, where data privacy and protection are provided?	
1	Security of the central data storage is not considered
2	Security of the central data storage is considered, but data can leave the storage for analytics purposes to uncontrolled environments
3	There is a controlled 'sandbox' environment where data from the shared data storage can be used for analytics purposes, but data from the central data storage can still be used in uncontrolled environments
4	There is a controlled 'sandbox' environment where data from the shared data storage can be used for analytics purposes
5	There is external governance on the security and privacy of data for all data between hospitals

Indicators for organization could be to ask the privacy officer how awareness is created on privacy issues.

Awareness	
1	There is little awareness of privacy issues
2	Privacy of patient data is discussed when employee is hired
3	There are documents available for education purposes on privacy of patient data
4	There is training available for education purposes on privacy of patient data
5	Privacy is on top of mind for all employees that are dealing with data

For the question on technology ask the head of business intelligence how and when in the process pseudonymisation is done.

Anonymization	
1	There is no technical solution for pseudonymization of data
2	There is a technical solution for pseudonymization of data but has to be executed through a manual process
3	Pseudonymization of all data can be done through an automatic process
4	Privacy by design is the used approach throughout the whole engineering process
5	Data from different care providers can be shared, while pseudonymization of this data is in place

Innovation

The questions on innovation are most likely to be answered by someone responsible for innovation within the academic hospital. The CIO and the person responsible for research ICT might offer answers to the questions.

Usage of big data	
You can use big data to prove theories or use to discover new theories or correlations. To prove a theory, you might not need big data after all.	
1	We do not use data analytics

2	We do use analytics, but nog on big data
3	We use big data analytics to prove suspected theories (retrospective)
4	We use big data analytics not only retrospective, but also proactive: the data gives us new insights and we can explore this data to come up with new theories
5	We use big data analytics not only retrospective, but also proactive and combine our data with other data sources such as weather data to discover unsuspected patterns

To know how actively the UMC participates in innovation on a national level, ask the CIO or innovation team what their role is in the national initiatives already existing such as the meetings set-up by the NFU (Nederlandse Federatie van Universitair Medische Centra)

The UMC actively **participates** in innovation	
1	Innovation is not part of the academic hospital
2	Innovation is part of certain departments of the academic hospital
3	Innovation is hospital-wide part of the academic hospital
4	Besides innovation being a part of the entire UMC, the UMC participates in innovation programs on a national level
5	Besides innovation being a part of the entire UMC, the UMC has a leading role in innovation programs on a national level

Multidisciplinary teams	
1	Big data analytics are not performed
2	Big data analytics are performed purely by individuals
3	Big data analytics are performed purely by individuals, but do share experiences with others in their team
4	Big data analytics are performed purely by individuals or teams, but share their knowledge and best practices horizontally across the organization
5	Big data analytics are performed by multidisciplinary teams

This question could best be answered by someone of the innovation team, if present.

Creativity: There is time to think creatively for the innovation team	
1	We do not have people that have a role involving innovation
2	There is no time for creativity for people involved with innovation in their day-to-day conduct of work
3	Some working hours could be spent on creative ideas, but this is not encouraged actively
4	Employees can spend time on creativity and this is encouraged actively by their manager
5	Employees can spend time on creativity and this is encouraged actively by their manager. Employees are educated on creative thinking techniques or 'design thinking'

MATURITY MODEL VERSION 2.1

The only changes that were made for version 2.1 were made to the instructions of the model and the attribute Anonymization was renamed to Pseudonymization.

INSTRUCTIONS

Instructions on how to use the maturity model will be discussed in this section. UMCs that want to assess their big data analytics maturity are advised to go through all steps described in this section.

Step 1: Role assignment

The hospital that wants to prepare for the assessment should first assign certain roles before it can start. One person should execute the assessment and define the employees that need to be involved, and one person high in the organization should sponsor the assessment.

Assessor

The assessor should execute the maturity assessment. The assessor needs to be familiar with big data and have enough experience within the hospital to know which persons to ask for the assessment. The assessor needs to collect data from IT and business. The assessor could be someone working on innovation or involved with research ICT. The assessor should have knowledge on all the different domains.

Sponsor

Since multiple people within the hospital will be asked to spend time on the assessment and the assessment might lead to change, someone high in the organization should sponsor the project. He or she should let the organization know that the project is important and that people should cooperate.

Step 2: Collecting the data

The next step is to start collecting the relevant data to answer the questions of the model. Roles that should be asked for input are (if present within the hospital):

Roles
IT architect
Head of Business Intelligence
Innovation manager
CIO
Privacy/ security officer
Researcher using big data analytics, or more advanced analytics in the hospital

The assessor should have enough knowledge of the hospital to know if these roles are present within the organization, and if not find suitable replacements.
Data can be collected by handing out the questions to the specific people chosen, or if there is not enough knowledge of big data analytics to let the questions directly be answered, interviews can also be sufficient.

Step 3: Determining the maturity

When enough data collected to answer all questions, the questions should be answered using the scoring form, which can be found in Appendix E. When in doubt, the assessor should decide if more information is needed, or if he/she can make a decision himself. After finishing answering all the questions, the visualizations of the current maturity should be made. The questions should only be answered by the assessor.

A guideline to help answering the questions can be found in Appendix B.

Step 4: Interpretation

The model is a tool to quickly assess and illustrate the current maturity in big data analytics within the hospital. This is purely descriptive, and does not provide immediate improvement potentials. Therefore, interpretation of the outcomes is needed. There are a couple of steps to interpret the results:

1. As-is landscape
   The assessor should start by analyzing the current landscape within the organization. Define what the current limitations are, and how big data analytics is currently placed within the organization.
2. To-be landscape
   A hospital that wants to assess their current big data analytics maturity, most likely has goals concerning this topic. Make the to-be maturity clear for the academic hospital. For the to-be maturity a period of two years is advised.
3. Gap analysis
   Once both the as-is and the to-be landscape are clear, the assessor should analyze the gap between these landscapes. These gaps should be connected to domains of the maturity model.
4. Recommendations
   The gaps, and the related domains, should be translated by the assessor to actionable recommendations. This does not mean that the hospital should immediately strive for the highest maturity level, but to a reachable higher maturity that is aligned with the strategy of the hospital.

Step 5: Validation
To ensure that the assessor collected the right information and interpreted the information in the right way, the assessor should make a report and/or presentation to all persons that were involved. An example of a report can be found in Appendix F. This should lead to a feedback session where the results might be slightly adjusted if needed. If a feedback session is difficult to organize, a survey could be sufficient. An example of the survey can be found in Appendix D.

After the validation, the report and/or presentation should be finalized.

1. Table of contents
2. Introduction
3. Explanation of the model
   a. Maturity levels
   b. Domains and attributes
4. Research method
5. Results
   a. Strategic alignment
   b. Governance
   c. …
6. Conclusions and recommendations

FIGURE 38 EXAMPLE TABLE OF CONTENTS FOR THE REPORT

Step 6: Distribution of report
The final step of the maturity assessment is distribution of a report with the findings of the maturity

assessment. Figure 10 shows a suggested table of contents for the report. After finishing the report, the assessor should actively spread the results within the UMC.

APPENDIX C: INTERVIEWS SETUP AND PRODUCT

During this research project three types of interviews were conducted. In this section an overview is provided of interviewees per type of interview and the main findings from these interviews. Transcripts are not provided because of privacy reasons. When necessary, the author can be contacted for these transcripts.

INTERVIEWS WITH EXPERTS

Questions that were asked were:

1. What is your expertise on big data analytics?
2. What is your career background?
3. What is big data?
4. How can big data bring value to healthcare?
5. Should academic hospitals use big data?
6. What is key when embedding big data analytics in an organization, specifically an academic hospital?
7. What challenges arise when embedding big data analytics in an organization, specifically an academic hospital?
8. How would you measure big data competencies in an academic hospital?

LIST OF INTERVIEWEES AND KEY FINDINGS

Expert 01 - Big *data master at Deloitte – 17 years of experience – Transcript number 01*
Expert 02 - *Team leader of the big data team at Deloitte – 11 years of experience – Transcript number 02*
Expert 03 - *Big data organizations expert at Deloitte – 14 years of experience – Transcript number 03*

DEFINITION OF BIG DATA

Expert 01: “Big data is defined by the three V’s: volume, variety and velocity. To be big data, it should adhere to all of those V’s.”
Expert 02: “Big data is defined by the three V’s: volume, variety and velocity. Two out of three should be present in order to call the data big data.”
Expert 03: “Big data is partly the three V’s: volume, variety and velocity. But is it mostly about what you do with the data. It should be used in an exploratory way to really be considered as big data analytics. It is about combining different sorts of data without knowing what you might find.”

HOW CAN BIG DATA BRING VALUE?

Expert 01: “You should see it as an investment to store the data you cannot store in a data warehouse. You should not know beforehand what to do with the data. Do not think with use cases in mind, then regular analytics approaches could be sufficient. Use the data exploratory.”
Expert 02: “An academic hospital should really understand what big data means and not go for it because it is a buzzterm. It will only bring value if used right.”
Expert 03: “Big data is as useless as small data when used wrong. Big data should be an addition to traditional business intelligence, not as a replacement. If you use it as a journey of discovery, it can bring value to an academic hospital.”

HOW TO MEASURE BIG DATA CAPABILITIES?

Expert 01: “Deloitte has an Insight Driven Organization framework that could serve as a basis.”
Expert 02: “I look at the Insight Driven Organization framework to determine if an organization is ready for big data. “

Expert 03: “I would use the Insight Driven Organization framework from Deloitte as a basis. But you should really look at more than data and information technology Strategy, process, value creation, people, stakeholders, risks and ethical questions have to be considered.”

CHALLENGES

Expert 01: “Privacy is very complicated when it comes to big data. The privacy aspect limits the big data scientist. This, and security of the data, is still in a beginning phase in big data tooling. With databases this is very advanced and developed, but this is the greatest challenge for implementing a big data environment.”
Expert 02: “Technically there are still issues that are tricky such as real-time data. Furthermore, ownership of data is very important because of the privacy laws and legislation. I belief that a person generating data is owner of that data.”
Expert 03: “Change management is very important. The people should accept new technology, or it will never work out for a company.”

OTHER

Expert 01: “Big data is believed to be very difficult, but just start small and don’t hesitate to experiment a little.”
Expert 03: “I believe that companies will be valued on their data in the future. I absolutely believe in the role of the Chief Data Officer. Data should be in the highest management.

INTERVIEWS ON NATIONAL INITIATIVES

During the research project, several board members from national initiatives were interviewed to get an understanding of the number of national initiatives on big data analytics, their strategies and the way they relate to other initiatives and healthcare providers.

The questions that were asked were:

1. Can you elaborate on the organization you work for?
2. What is the goal of this organization?
3. Who funds the organization?
4. Who determines the strategy of the organization?
5. What is the strategy of the organization for the next years?
6. How does the organization relate to other national initiatives and healthcare providers?

LIST OF INTERVIEWEES AND FINDINGS

In the overview below, all initiatives that were interviewed are shown including the role of the interviewee and the aim of the initiative. A transcript number is provided as a reference number to request the transcript.

Transcript number	**Initiative**	**Interviewee**	**Aim of the initiative**
04	Parelsnoer	Coördinator ICT	National database for diseases
05	NFU AcZie	Chairman	Meeting of all CIO’s from the academic hospitals
06	NFU project ‘Registration at the source’	Project leader	NFU project for registering data once, and then using it multiple ways

07	Sleutelnet	Project coordinator	A collaboration between healthcare providers in the area of Zuid-Holland on digital communication in healthcare
08	European Science Cloud	Chair of the High Level Expert Group	European initiative towards open science data for research FAIR data protocol – Initiator of the FAIR Data Principle – The FAIR Data Principles is a guideline for those wishing to enhance the reusability of their data holdings
09	Personal Health Train	Initiator	The Personal Health Train connects health data and makes it more usable by bringing the research to the data and not by centralizing the data
10	Informatieberaad	Initiator	Healthcare professionals and the Ministry of healthcare contribute to a sustainable information system

SOUNDING GROUP SESSION ON THE FUTURE OF BIG DATA IN HEALTHCARE

The sounding group on big data in healthcare consisted of three LUMC employees.

Visionary 01 – Counselor on ICT innovation – 20 years of experience
Visionary 02 – ICT architect and counselor – 20 years of experience
Visionary 03 – Counselor on ICT innovation – 2 years of experience

INTERVIEWS FOR THE MATURITY ASSESSMENTS

During the three case studies at Radboud UMC, VUMC and LUMC several interviews were conducted. For each of these case studies, a list of interviewees is provided including their role in the organization and the domain of the model their answers were used for.

Academic hospital	**Transcript reference number**	**Role in the organization**	**Domain**
Radboud UMC	10	Technical physician	Strategic alignment, Innovation, Organization
Radboud UMC	11	Research ICT architect	Strategic alignment, Information

			Technology, Data, Organization
Radboud UMC	12	Program manager HereIsMyData	Information Technology, Data
Radboud UMC	13	Director of REshape	Strategic alignment
Radboud UMC	14	Manages Business Intelligence & Analytics	Strategic alignment, Governance, Data, Privacy, Organization, Innovation
Radboud UMC	15	Head of research group in Radiology	Organization, Innovation
VUMC	16,17	ICT manager research, education and innovation	All
VUMC	16,17	Information manager research	All
VUMC	18	Head of Business Intelligence	All
LUMC	19	Innovation counselor and Research ICT	Strategic alignment, Information Technology, Organization, Innovation
LUMC	20	CIO	Strategic alignment, Organization, Innovation
LUMC	21	CISO	Privacy, Innovation
LUMC	22	Head of integration and development	Strategic alignment, Data, Information Technology, Organization, Privacy
LUMC	23	Privacy Officer	Privacy
LUMC	24	Information manager research	Data, Information Technology, Privacy
LUMC	25	Manager Information Management and Architecture	Strategic alignment, Information Technology, Organization, Innovation

DOMAIN EXPERTS

For the domain privacy and innovation two experts were conducted by phone. The main takeaways are discussed.

Domain expert 01 – Expert on innovation culture in academic hospitals – 5 years of experience

“Academic hospitals are not fast enough in adaption to disruptive technology because of their culture. You should create an incubator to operate faster than the main organization, because otherwise innovation will be too slow. This way, projects can also fail because they are small and do not immediately disrupt the hospital. Failure should be possible, because innovation has unexpected results. By placing innovation directly in an organization, it is like trying to make a square out of a circle: it will not work.

Conditions for innovation are a creative mindset that can be learned by employees. Furthermore, time, money and diversity are the power behind innovation. Ideas for innovation should not be sought within the academic hospital, but in the outside worlds.

There is a difference between innovation and improvement. With improvement you look at the existing organization and make improvements, while with innovation you can also look at something totally new and crazy. “

Domain expert 02 – Expert on privacy laws – 21 years of experience

“Anonymization and pseudonymization should be in place in all academic hospitals. Privacy by design should be standard for all UMCs. Privacy is guaranteed by having strong IT facilities but also good protocols. For the healthcare industry, informed consent is necessary. The patient must agree that his data is used for research practices.

With the new privacy laws and legislation it will be very difficult for academic hospitals to do scientific research as they do this now because they are not treating data with the right amount of care. You should be open on your policies and information technology to secure data and guarantee privacy of patients. “

APPENDIX D: SURVEY SETUP AND PRODUCTS

Two types of surveys were used for this research project. The first type of survey is to validate the construct of the model and the second type of survey is to validate the obtained results during a case study.

SURVEY QUESTIONS TO VALIDATE CONSTRUCT OF THE MODEL

The following survey questions, based on the evaluation template from Salah et al. are shown below.

MULTIPLE CHOICE QUESTIONS

The following multiple choice questions were stated in the survey. Each of these questions had five possible answers, using the Likert scale (Wikipedia, 2017). These were strongly disagree, slightly disagree, neither disagree or agree, slightly agree and strongly agree.

Maturity levels

1. The maturity levels are sufficient to represent all maturation stages of the domain (Sufficiency)
2. There is no overlap detected between descriptions of maturity levels (Accuracy)

Domains and attributes

3. The domains and attributes are relevant to the domain (Relevance)
4. Domains and attributes cover all aspects impacting/involved in the domains (Comprehensiveness)
5. Domains and attributes are clearly distinct (Mutual exclusion)
6. Domains and attributes are correctly assigned to their respective maturity level (Accuracy)

Maturity model

Understandability

7. The maturity levels are understandable
8. The assessment guidelines are understandable
9. The documentation is understandable
10. The domains and attributes are correctly assigned in their maturity level

Ease of use

11. The scoring scheme is easy to use
12. The assessment guidelines are easy to use
13. The documentation is easy to use

Usefulness and Practicality

14. The maturity model is useful conducting assessments
15. The maturity model is practical for use in industry

OPEN QUESTIONS

The following ten open questions were stated in the survey.

Q1. Would you add any maturity levels? If so please explain what and why?
Q2. Would you update the maturity level description? If so please explain what and why?
Q3. Would you add any domains or attributes? If so please explain what and why?
Q4. Would you remove any of the domains or attributes? If so please explain what and why?
Q5. Would you redefine/update any of the domains or attributes? If so please explain what and why?
Q6. Would you suggest any updates or improvements related to the scoring scheme? If so please explain what and why?

Q7. Would you suggest any updates or improvement related to the assessment guidelines? If so please explain what and why?
Q8. Would you like to elaborate on any of you answers?
Q9. Could the model be made more useful? How?
Q10. Could the model be made more practical? How?

SURVEY QUESTIONS TO VALIDATE RESULTS OF MATURITY ASSESSMENT

For each of the questions from the maturity model, the answers that determine the current maturity and the to-be maturity of the academic hospital are validated. For every domain, the following questions are stated:

Please pick the best fitting answer for each question on the current maturity. If you agree with the currently stated answer, choose that answer.
Mark only one oval per row

For each of the four attributes of that domain, one answer out of the five possibilities has to be chosen. An example is shown in Figure 39.

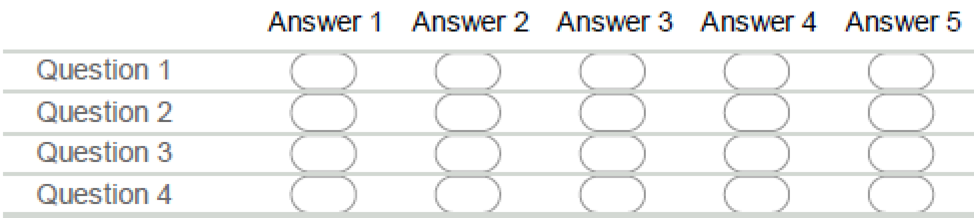


FIGURE 39 EXAMPLE OF SURVEY QUESTION TO VALIDATE A MATURITY ASSESSMENT

Similarly, a question on the to-be maturity is stated for each domain:

Please pick the best fitting answer for each question on the to-be maturity in two years' time. If you agree with the current stated answer, choose that answer.
Mark only one oval per row.

The answers to this question are similar to the previous question, visible in Figure 39.

At the end of each domain, two open questions are stated:

1. If there is anything you want to add for the current maturity, please elaborate:
2. If there is anything you want to add for the to-be maturity, please elaborate:

These questions are repeated for each of the seven domains.

RESULTS OF THE SURVEYS

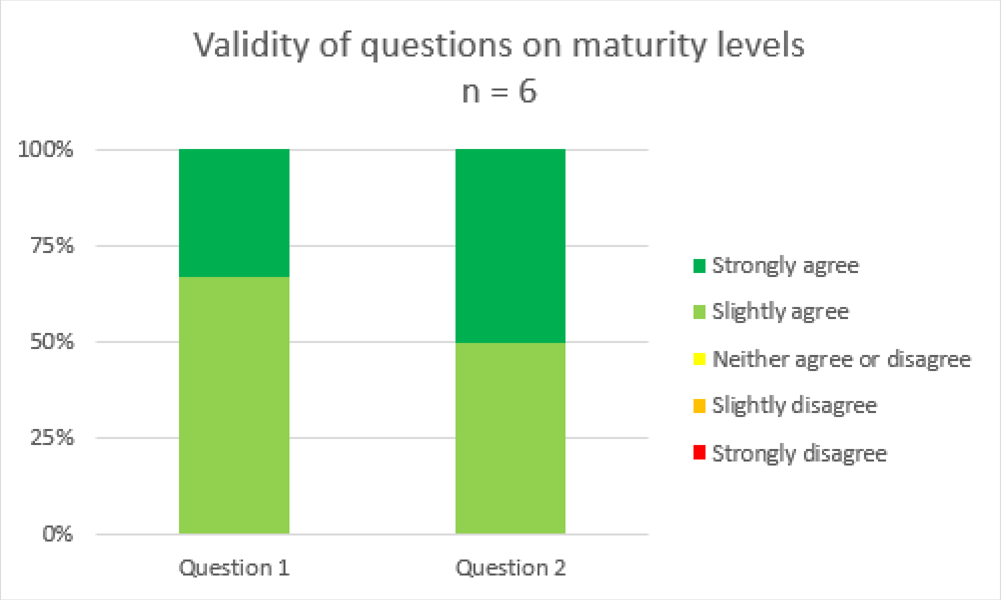
In this section the results of the surveys are provided. Firstly, the responses to the surveys on the validity of maturity model version 1.0 are shown. Secondly, the responses to the surveys to validate the case studies are shown for first Radboud UMC and then VUMC. Thirdly, the responses to the surveys on the validity of maturity model version 2.0 are shown and finally, the responses to the surveys to validate the case study at LUMC are shown.

VALIDITY MATURITY MODEL VERSION 1.0

The survey on the validity of the maturity model version 1.0 was answered by six respondents. Three were expertson the matter, two were from the VUMC and one was from the Radboud UMC.

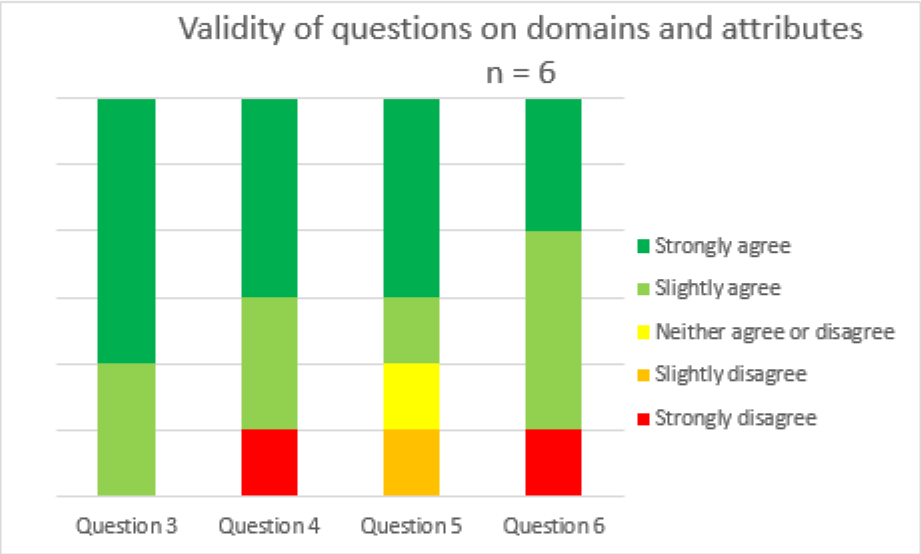
Question 1 to 2

Q1 Maturity levels [The maturity levels are sufficient to represent all maturation stages of the domain (Sufficiency)]
Q2 Maturity levels [There is no overlap detected between description of maturity levels (Accuracy)]



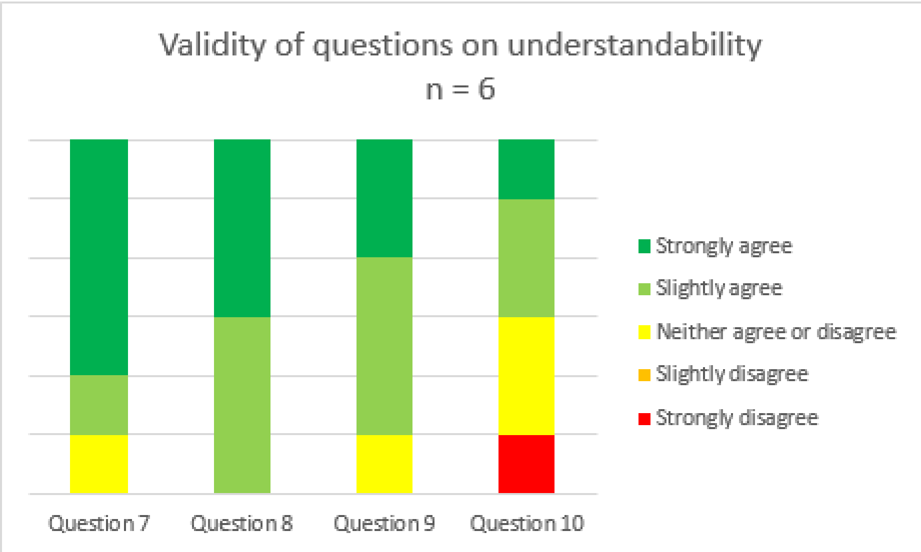
Question 3 to 6

Q3 Domains and attributes [The domains and attributes are relevant to the domain (Relevance)]
Q4 Domains and attributes [The domains and attributes cover all aspects impacting/involved in the domain]
Q5 Domains and attributes [The domains and attributes are clearly distinct (Mutual Exclusion)]
Q6 Domains and attributes [The domains and attributes are correctly assigned in their maturity level (Accuracy)]



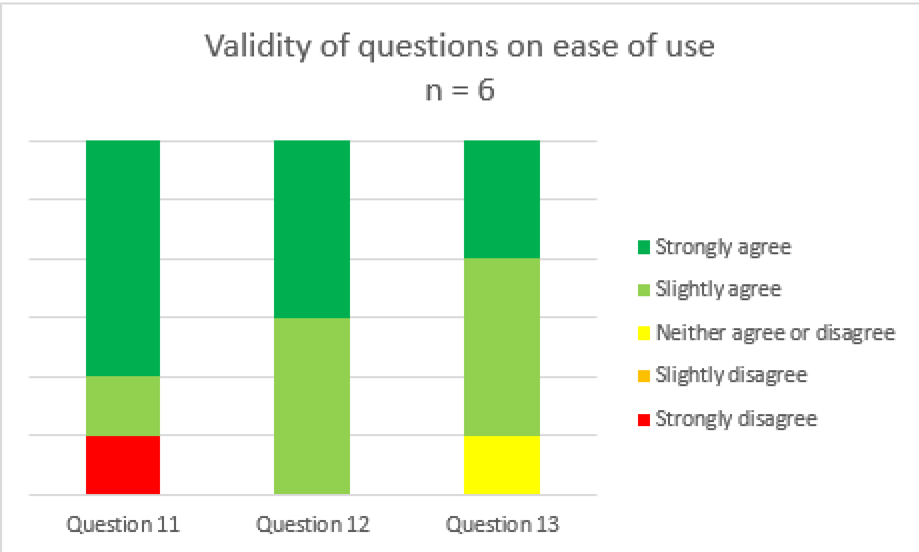
Questions 7 – 10

Q7 Understandibility [The maturity levels are understandable]
Q8 Understandibility [The assessment guidelines are understandable]
Q9 Understandibility [The documentation is understandable]
Q10 Understandibility [The domains and attributes are correctly assigned in their maturity level (Accuracy)]



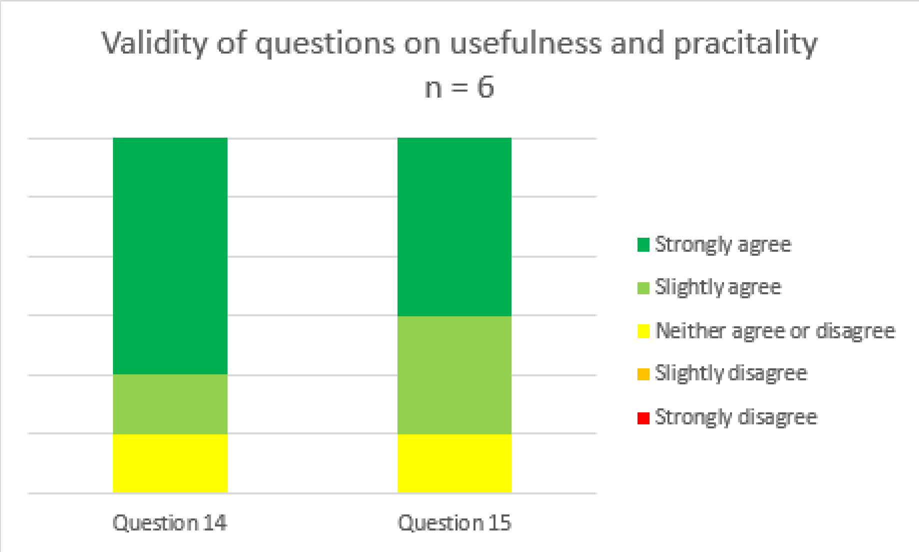
Question 11 – 13

Q11 Ease of use [The scoring scheme is easy to use]
Q12 Ease of use [The assessment guidelines are easy to use]
Q13 Ease of use [The documentation is easy to use]



Questions 14 - 15

Q14 Usefulness and practicality [The maturity model is useful conducting assessments]
Q15 Usefulness and practicality [The maturity model is practical for use in industry]



Would you add any maturity levels? If so please explain what and why?
No, five times
Define/describe big data analytics

Would you update the maturity level description? If so, please explain what and why?
No, four times
Sometimes you write about protocols and sometimes about processes. For me it is also not clear what the difference is between academic hospital and national level. In optimized you write about 'discover new insights' en at IT it is about 'generate knowledge'
Bij sommige vragen is het duidelijk geen 3 maar valt de keuze tussen 2 en 4

Would you add any domains or attributes? If so please explain what and why?
No, six times

Would you remove any domains or attributes? If so please explain what and why?
No, six times

Would you redefine/update any of the domains or attributes? If so please explain what and why?
No, two times
The Privacy domain has some attributes with the same name as some other domains (Organization, Technology, Governance). So this means some overlap and it is also confusing. Also with the Stewards attribute in the Organization domain. Data Stewardship is something that often is described as part of Data Governance.
One of the characteristics of Big Data analytics is that it is not about ETL but about ELT. It is also not only about Data Ware Houses which is an old (but still relevent) BI concept. Regarding Data one opportunity is the use of external datasources. So not only data which is generated by the hospital. One top priority regarding privacy is Security. That topic is not mentioned anywhere. Culture is a big topic on its own. There is a lot written about culture and there can be a lot interpretations.
Minder nadruk op IT. Datamanagement door de business is niet slechter dan door IT
Misschien ook nog ergens overheids/maatschappelijke visie/stimulatie/druk erin verwerken

Would you suggest any updates or improvements related to the scoring scheme? If so please explain what and why?

No, two times
It would be easier if it would be possible to click on a answer instead of filling in the number of the answer.
Your domains consist of Governance, Technology and Organizations. Those are also Attributes within the domain of Privacy. Atribute Data Usage is part the domain Innovation where the atribute Usage is also part of the domain Data.
Bij sommige vragen is het duidelijk geen 3 maar valt de keuze tussen 2 en 4
CIO (or someone comparable) maybe replace it for C-Level or CxO

Would you like to elaborate on any of you answers?

No, five times
CMMI good model for maturity. All dimensions are relevant. Difference per level on the dimension is debateable but needs to be tested.

Could the model be made more useful? How?

No, 4 times
Nee, current model is very useful
Ik denk dat het een hoog kennisniveau vereist om dit model te gebruiken, maar wellicht is dit onoverkoombaar en kan je het niet afnemen bij elke werknemer.

Could the model be made more practical? How?

No, two times
Very practical
I would not use the excel with all the different tabs and dropdowns.
Works well
De uitleg van het model staat behalve in de documentatie nu ook kort in rapport. Misschien drie versies: Documentatie + rapport (zonder uitleg) of enkel rapport met korte uitleg. #muggeziften

Do you have any other feedback regarding the model or this questionnaire?

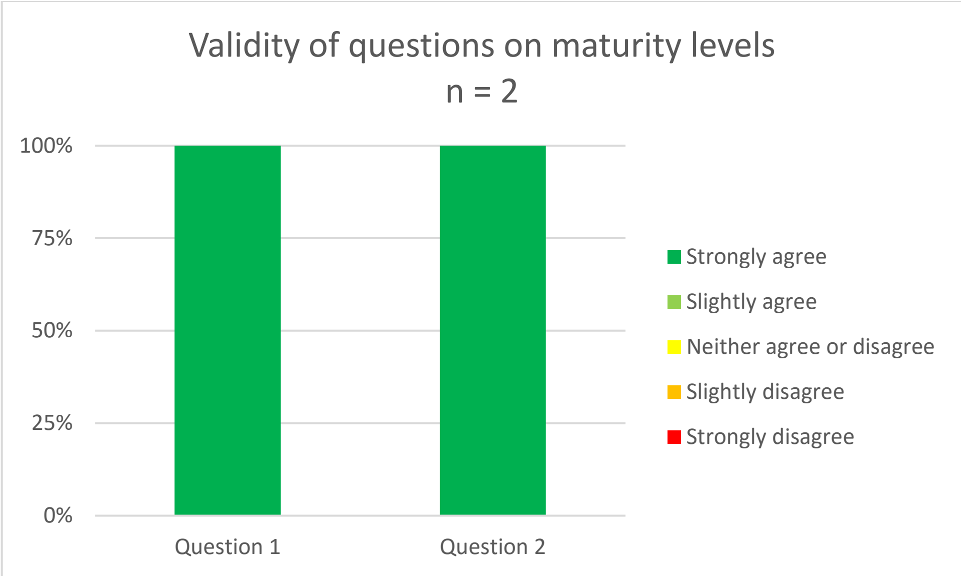
No, four times
Some questions does not cover 'Big Data', but more the maturity in general (for example an hospital can have a very mature IT organization, Data Governance policies, etc. but not do anything with Big Data). Especially the domains Governance and Privacy are not very focused on Big Data.
Good luck

VALIDITY MATURITY MODEL VERSION 2.0

The survey on the validity of the maturity model version 1.0 was answered by two respondents. The respondents were interviewees from the LUMC.

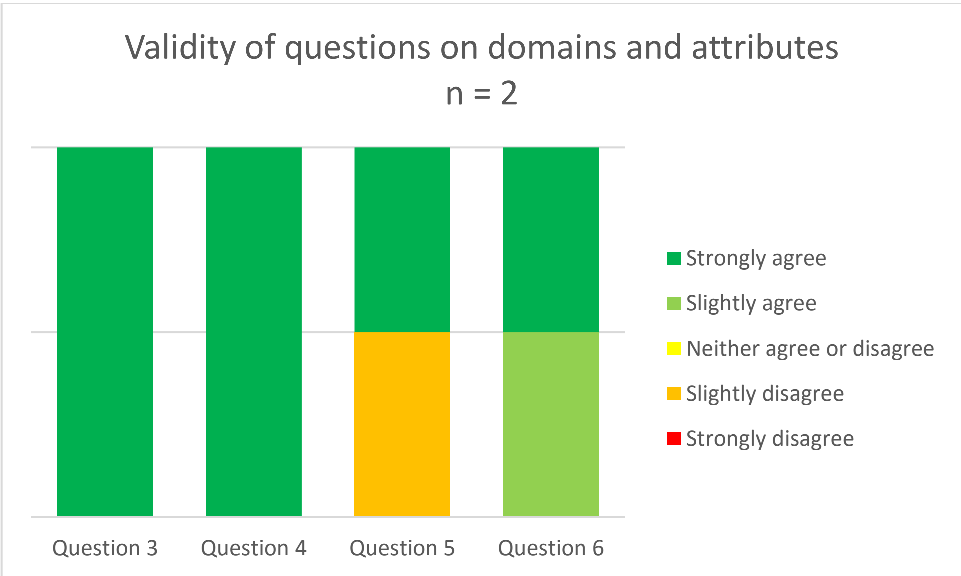
Question 1 to 2

Q1 Maturity levels [The maturity levels are sufficient to represent all maturation stages of the domain (Sufficiency)]
Q2 Maturity levels [There is no overlap detected between description of maturity levels (Accuracy)]



Question 3 to 6

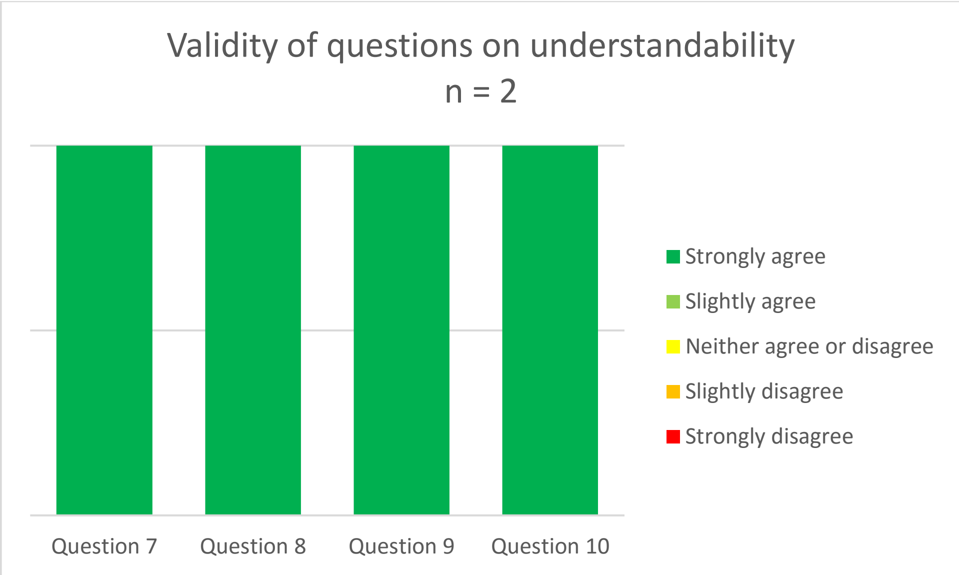
Q3 Domains and attributes [The domains and attributes are relevant to the domain (Relevance)]
Q4 Domains and attributes [The domains and attributes cover all aspects impacting/involved in the domain]
Q5 Domains and attributes [The domains and attributes are clearly distinct (Mutual Exclusion)]
Q6 Domains and attributes [The domains and attributes are correctly assigned in their maturity level (Accuracy)]



Questions 7 – 10

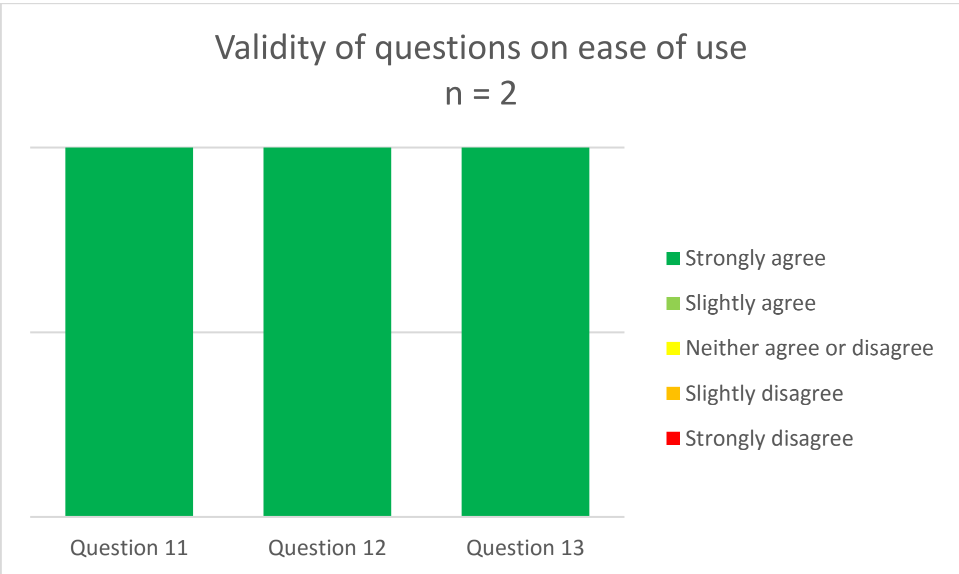
Q7 Understandibility [The maturity levels are understandable]
Q8 Understandibility [The assessment guidelines are understandable]
Q9 Understandibility [The documentation is understandable]

Q10 Understandibility [The domains and attributes are correctly assigned in their maturity level (Accuracy)]



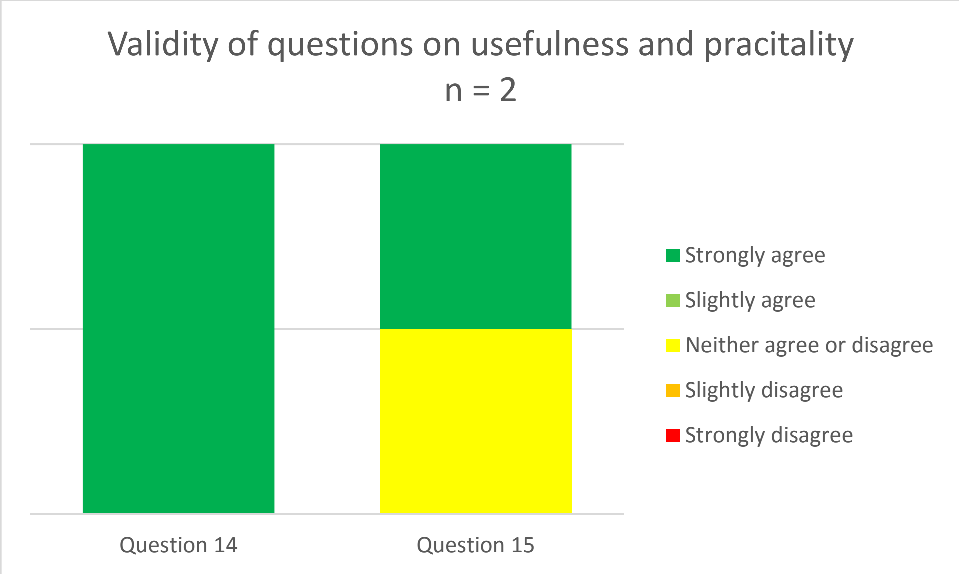
Question 11 – 13

Q11 Ease of use [The scoring scheme is easy to use]
Q12 Ease of use [The assessment guidelines are easy to use]
Q13 Ease of use [The documentation is easy to use]



Questions 14 - 15

Q14 Usefulness and practicality [The maturity model is useful conducting assessments]
Q15 Usefulness and practicality [The maturity model is practical for use in industry]



Would you add any maturity levels? If so please explain what and why?
Nee
Ik vroeg me af of je iets zou moeten vinden van machine- / deeplearning. Gaan we modellen integreren?

Would you update the maturity level description? If so, please explain what and why?
Nee, dit komt overeen met bestaande modellen en is hiermee direct begrijpbaar.
-

Would you add any domains or attributes? If so please explain what and why?
Nope
-

Would you remove any domains or attributes? If so please explain what and why?
Ik vond het domein innovatie wat lastig. Het gaat over bigdata. wat nu een innovatief onderwerp is. De toegevoegde waarde van dat domein snapte ik niet.
-

Would you redefine/update any of the domains or attributes? If so please explain what and why?
Soms kon ik mijn antwoord niet zo makkelijke kwijt. Die heb ik aangeven.
-

Would you suggest any updates or improvements related to the scoring scheme? If so please explain what and why?
nee
-

Would you like to elaborate on any of you answers?
Heb ik direct gedaan
-

Could the model be made more useful? How?
Nee
Ik vond de termijn van 2 jaar nog wat beperkt. Daarmee wel overzichtelijk overigens. Zou het niet nuttig zijn om ook aan te kunnen geven wat je vindt dat je stip op de horizon is. Voor bijvoorbeeld 5 jaar.

Could the model be made more practical? How?
Nee
-

APPENDIX E: SCORING FORM

The scoring form is attached below.

Big data analytics
maturity model for D

APPENDIX F: REPORTS ON MATURITY ASSESSMENTS

The reports are confidential.