Universiteit Leiden

ICT in Business

"Conceptualization and validation of a set of metrics to assist Data Driven Decision Making in an Agile setup. A case study of KLM - Royal Dutch Airlines"

Name: Konstantinos Tsogkas
Student-no: s1580752

Date: 20/06/2017

1st supervisor: Dr. Hans Le Fever (LIACS)
2nd supervisor: Dr. Arno Knobbe (LIACS)

MASTER'S THESIS

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

Acknowledgements

My deepest gratitude goes to the team from KLM – Royal Dutch Airlines, that involved me into their project and granted me the opportunity to fulfill the necessary steps for my data collection. A big thank you to Robin Spierings, Elke Douma and Renee Dutscher for their assistance and their motivation. I would also like to thank my manager during my internship within the company, Joost Olieroock, who guided me and supported my choices while I carried out my internship within the department of Digital Analytics.

A big thank you goes to Harry Fulgencio, who willingly assisted me during my journey, guided my thoughts and contributed a big deal to the final form of this research thesis. I would also like to thank my first supervisor, Dr. Hans Le Fever, who patiently provided his feedback during the stages of my work, as well as my second supervisor, Dr. Arno Knobbe.

Last but definitely not least, I would like to thank my family; Athanasios, Stavroula and Evangellos. Your support through these years cannot be described in a few sentences. I will always love you and I will always be grateful for what you have done for me.

Enjoy,
Konstantinos.

“Επιστήμη ποιητική ευδαιμονίας”
Πλάτων, 427-347 π.Χ.

“Knowledge generates prosperity”
Plato, 427-347 B.C.

Management Summary

This particular research study aims to inform the reader about the factors that contribute to the decision making of certain personas in a software delivery setup. Data driven decision making is a science that shows little application in the organizational level, a fact that should be taken under consideration, as the industry proves that the value of informed decisions is increasing nowadays more than ever. Decision makers seek new ways to inform their reasoning, and they regard the knowledge that derives from data priceless, as it may be used as a significant competititve advantage.

KLM **-** Royal Dutch Airlines is one of the many large scale organizations that wants to exploit all kind of data in its' possession. Much research has been conducted in non-profit or educational organizations, but in literature only few automated ways are described, that may assist a company to be more data driven. A framework that enlists a number of metrics under certain categories will be presented here as a method for scoring drivers that affect decisions, which eventually assists the prioritization of software requirements in an Agile setup. Applications of this framework with different decision makers will be shown, and the results of this procedure will help us understand in which ways we can improve this framework in the future.

Very few measurable metrics can be used within the organization in order to frame a holistic approach about a software requirement. It is interesting to see that Product Owners consider several metrics to be way more important than others regarding the delivery of their software requirements. Our proposed framework could be of use to them, if they had proper access to information related to the metrics listed and it will be of utmost value to provide a document in a form of a sheet, that will be able to inform all of the stakeholders the reasons why several items are delivered over others. Metrics within the bracket of business value and user value are appreciated more when it comes to decide upon a requirement, and if Product Owners mean to orient their decisions in a broader perspective, they need to have access to metrics that are beyond rough data. Moreover, each decision maker perceives a metric as a quantifiable measurement, that its value may be used for numerical comparison; not much attention is given to the metrics that cannot be educated through quantitative data. Finally, Product Owners appreciate the possibility to evaluate more metrics, but they are limited due to the fact that they cannot find relevant information about them.

We must take into account that the conclusions to which we arrive apply for the present case study. It is only a wishful thinking that this framework shall be applied in different organizations in the future, with similar organizational setups and with common targets. The communication amongst stakeholders regarding software delivery is of utmost value, and only with further investigation and larger scale implementation shall we be able to generalize our findings and reinforce the validity of the current framework.

Keywords: Data Driven Decision Making, Data Science, Prioritization Methodology and Agile Framework.

Table of Contents

1.Introduction

1.1 Background

Digitization of products is increasing in many organizations; a need has been created for managers to monitor their operational performance, and gain insights through their customer's online journey and their historical data. As stated by Power (2013), "today data-driven decision support is used for a variety of purposes including operational and strategic business intelligence queries, static and real-time performance monitoring, and customer relationship management".

As more and more organizations tend to adopt data-driven decision management in their culture, the quality of data gathered and the interpretation of these datasets comprises the effectiveness of this data-driven approach. A study from the MIT Center for Digital Business found that organizations driven most by data-based decision making (DDDM) had 4% higher productivity rates and 6% higher profits.

Good examples of DDDM applications can be observed in organizational teams working within Digital Departments. Decision makers can be recognized also in the faces of people working in agile environments in different frameworks regarding software component delivery; it is an environment that creates and supports a culture that encourages a team of people to work towards common goals. It is accomplished by incorporating the importance of individuals and their interactions, in order to achieve quality and acceptance of frequent change in the company culture[1]. A Scrum Master or a Product Owner (PO), as recognized in a Scrum process framework for agile development, are decision maker personas that are responsible for making the right decisions which align with their organization's goals. Their decisions have to be based on the information that they acquire from their customer's data, in order to deliver their products faster, tailored to their customer's needs.

But what happens when they are required to prioritize a number of requirements that address to multiple stakeholders and that are time demanding in the same time? How can a decision maker enhance his decisions to be more data driven and make the prioritization of the items' delivery transparent to his stakeholders?

Many drivers may influence a decision maker; revenue, business value, risk reduction, are considerations that may assist the steering of his choices. But is there a standardized way to concede to the most effective evaluation of these drivers, minimizing the subjectivity involved in this decision making process, which is mostly done based on the gut feeling of the project leaders?

Managers and decision makers nowadays have the inquiry of how to interpret the data-driven concept and, moreover, how to apply it into their business. A need to recognize a pattern then appears, which will assist them to orient themselves that way. This pattern shall unveil itself within the scope of this research, applied in a large organization following initial interviews, conceptualization and operationalization of a set of metrics and, finally, quantitative and

[1] http://study.com/academy/lesson/what-is-an-agile-environment-definition.html

qualitative analysis of the behavior of decision makers, in order to reach to proposed optimizations.

1.2 Research Objective

The aim of this research is to contribute to the decision making process of decision makers identified as Product Owners within the Agile Scrum Framework, providing them a more data-driven methodology in order to prioritize the items that need to be delivered by their software teams. Through this research, we will identify which are the main drivers that inspire Product Owners to base their decisions upon through series of semi-structured interviews. In that sense, we will be able to identify a set of metrics, based on already existing data, that may be used in order to evaluate the significance of delivery of an item, in terms of software implementation.

After these metrics have been identified, we will implement a framework that is comprised of all the identified drivers that influence a decision. With this framework, we will work on the lists (Product Backlogs) that help Product Owners prioritize their items and we will actually try to score levels of delivery importance for each item. In that sense, we will be able to make the decision making process and, thus, the prioritization of the lists, a more transparent procedure for the decision makers themselves as well as for their stakeholders.

1.3 Research Idea

As Light et al. (2006) state, "the research on data systems and tools to support instructional decisions is a young and emerging field." According to Cromey (2000), there are technical and usability issues that occur when an effort is made towards the support of instructional planning; data storage, data entry, analysis, interpretation of data and the relationship between data and instructional practices. The growing interest in data-driven decision making tools is no doubt a direct response to these mounting pressures (Stringfield et al., 2005).

Several conceptual frameworks navigated towards this direction have already been introduced. One of them is presented in the research paper of Light et al. (2006), founded on the notion of what it means to be data-driven. Although that framework is referring to levels of organization across the educational system, it assumes that "individuals, regardless of where they are within a school system, have questions, issues, or problems for which data must be collected, analyzed, and examined in order to make informed decisions". An interesting question then would be to recognize which steps are followed by decision makers in a research setup dealing with delivery of software requirements.

Towards that path, it would be interesting to investigate if the decision makers are led by data to justify their decisions. Is DDDM part of the Product Owners cognitive skills, or are they leading their decision making process based on their experience and gut feelings? Tingling & Brydon (2010) state that "data-driven decision making capability is defined as the abilities of an organization to utilize data, information, and insight assets in a series of coordinated decision making processes in order to support, inform, or make decisions. The definition summarizes three different roles of data, information, and insight assets in decision making, which are to make a decision, inform a decision, and support a decision."

The goal of this research is to enhance the procedures involved in the first level of the data driven decision making framework under study, which refers to the acquirement and analysis of data. Decision makers must be aware which specific metrics can be used that may influence their decisions and to be taken under consideration in order to deliver a software requirement over another.

1.4 Research Questions (RQ)

For the formulation of the main RQ, considering the above, we conclude to the following format:

“Conceptualization and validation of a set of metrics to assist Data Driven Decision Making in an Agile setup. A case study of KLM - Royal Dutch Airlines”

The following research questions come up and need to be answered through this thesis work :

- ➔ What are the main criteria taken under consideration currently from decision makers in order to take decisions? How do Product Owners perceive them and how do they measure them?
- ➔ Can the prioritization of software requirements be more data driven? How can the education of this knowledge be enhanced?

1.5 Research Scope

The present research thesis is conducted within the organization of KLM Royal Dutch Airlines, in cooperation with the Excellence and Performance Department at KLM Headquarters, the Netherlands. This study is structured in stages of interviews that are three-folded.

In the first stage, through initial interviews we shall have an overview of the processing that is followed by Product Owners, in order to deliver a prioritized list of their software requirements that need to be implemented. The setup of all teams that are under examination is unified; all teams have adopted Agile Scrum methodology of working, with Product Owners being the decision makers in our case. Product Owners have the responsibility, amongst others, to refine their Scrum Product Backlog in collaboration with their stakeholders, or representatives, in order to deliver all the software requirements in the most time effective way.

The second stage is focused on their data driven decision making capabilities and based on the feedback that we will acquire from the first stage. Definitions of the drivers and the criteria that have been recognized shall be provided to them. At this stage, we will try to understand how do Product Owners perceive Data Driven Decision Making. Most conspicuously, we shall present to them all the metrics and drivers that we have recognized through the initial interviews and ask them whether they have sufficient data to understand them thoroughly. Similarly, with the same questionnaire, we will be able to realize in a quantitative manner which criteria are mostly evaluated by the POs while they prioritize. This will provide insights on which criteria are more important to focus on, and for which of those we can receive complete and understandable datasets that may provide information, which can lead to decisions. In that sense, the realization of these main drivers is a step closer to realize what kind of data is needed to provide them with better insights and enhance their decision making with the use of proper data.

In the third stage, we interview the POs with their Scrum Product Backlogs in order to apply these requirements on their decision making process. The scoring of importance for each metric is based on the Fibonacci sequence, and it depicts the levels of importance for each driver. Through this process, a median average shall come up, in order to reveal which items need to be delivered before others, and if this complies with their current prioritized list.

1.6 Thesis overview

In the first chapter we introduce the reader to the concept of this research study, providing information about the research objective, research idea, research questions and eventually research scope.

In the second chapter we present an elaborate literature review of the study, in order to familiarize the reader with the concept of Data Driven Decision Making, the applications of DDDM in large or small, for-profit or non-profit organizations, and the literature gap that we identify behind certain levels of identified frameworks, as well as our contribution to the optimization of this gap. Definitions of the identified metrics are also provided.

In the third chapter, we state the conceptualization of our research methodology and the design that we followed in order to obtain our qualitative and quantitative data.

In the fourth chapter, we outline the results and the outcomes of the three different stages of semi-structured interviews.

In the fifth chapter we proceed to the validation of the conceived idea, thus to the validation of the proposed set of metrics.

In the sixth chapter we answer the research questions by combining the results from both literature and the sets of qualitative and quantitative data that we have gathered. We also provide conclusions and recommendations of how a for-profit organization may be more data-driven and present ideas for further studies.

In the seventh chapter the references from the literature are being listed alphabetically and finally, in the eighth and last chapter we have a complete depiction of the appendices.

2. Literature Review

In order to understand which kind of bibliography hinders behind the vast topic of Data Driven Decision Making (DDDM), systematic literature review was conducted in order to identify existing frameworks, definitions and applications in real life cases. Related areas such as Decision Making and Requirements Prioritization were also investigated in order to obtain a thorough view of the subject. The major keywords that were used were: Data Driven Decision Making, Data Driven, Data Science, Big Data, Data Analytics, Decision Making, Requirements Prioritization, Prioritization Methods, Decision in Agile Frameworks, Key Performance Indicators, Metrics in Agile.

Besides keyword search, backward search was also implemented; in that sense, backward references search and backward author search was conducted, elaborating on the most chronologically recent papers. Along with the academic papers that were studied and referenced, which mostly came out of combinations of these aforementioned keywords in Google scholar and the database of the Leiden University Library, grey literature was also taken under consideration in order to formulate a holistic view of the topic. All of the references that were used are listed in alphabetical order in the Bibliography - References section.

2.1 Data Driven Decision Making (DDDM)

Data science is definitely the new discipline that will play a significant role in the industry for the years to come. Companies need to hire data scientists as today they are akin to the Wall Street "quants" of the 1980's and 1990's (Davenport, Patil, 2012). Data Science is often correlated to Big Data and Data Analytics. In this sense, a brief description of these aforementioned terms would assist us to outline their differences:

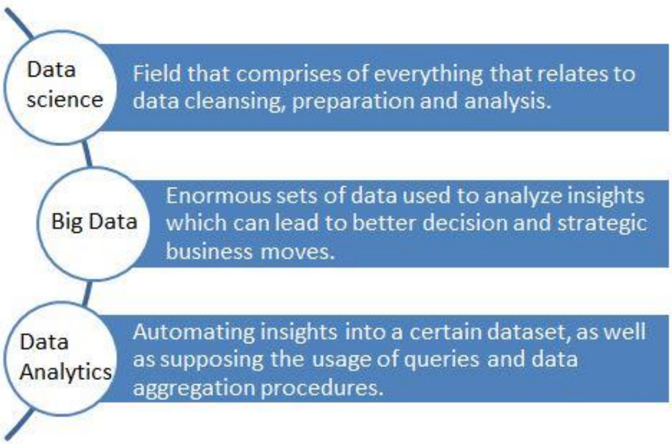

Figure 1: Description of terms

To get a better understanding of Big Data in particular, it refers to sets of data that are too large for traditional data processing systems and that require new technologies. Economist Prasanna Tambe, of New York University's Stern School, has examined the extent to which the utilization of Big Data technologies seems to help firms. He finds that, after controlling for various possible confounding factors, the use of big data technologies correlates with significant additional productivity growth. Specifically, one standard deviation higher utilization of big data technologies is associated with 1–3% higher productivity than the average firm; one standard deviation lower in terms of big data utilization is associated with 1–3% lower productivity (Provost and Faucett, 2013). "Data science is the connective tissue between data-processing technologies and data driven decision making", while in a higher level "data science is a set of fundamental principles that support and guide the principled extraction of information and knowledge from data" (Provost and Faucett, 2013).

There is a set of well-studied, fundamental concepts underlying the principled extraction of knowledge from data, with both theoretical and empirical backing. These fundamental concepts of data science are drawn from many fields that study data analytics.

- Extracting useful knowledge from data to solve business problems can be treated systematically by following a process with reasonably well defined stages.
- Evaluating data science results requires careful consideration of the context in which they will be used.
- The relationship between the business problem and the analytics solution often can be decomposed into tractable sub-problems via the framework of analyzing expected value.
- Information technology can be used to find informative data items from within a large body of data.
- If you look too hard at a set of data you will find something, but it might not generalize beyond the data that you are observing, overfitting a dataset.
  (Provost and Faucett, 2013)

In an effort to distinguish conventional decision making processes from data driven ones, we would realize that the endpoint is common for both, but the way of rationalizing the decision is different. In a sense that "Data Driven Decision Making (DDDM) is focused on building knowledge, and using a wide variety of data to construct knowledge, to better understand what is actually going on versus what is assumed" (policy data driven decisions), DDDM utilizes data in order to inform conventional decision making. It normally consists a principled technique to rationalize variable observations through the breakdown of data, rather than basing the decision on intuition.

Not much relevant literature studies relate to concepts of data driven decision making that find implementations beyond the educational sector. Much past literature on data-driven decision making has not contributed in a practical manner to help organizations build routines of data-driven decision making (Garvin, 2013). The gap between the research that has been implemented up until this point and this research scope, is that we shall try to operationalize the existing framework of Jia et al. (2015) according to the feedback that we received after the interviews with the decision makers under the case study within KLM - Royal Dutch Airlines.

According to Hall et al. (2015), the findings are not systematically integrated, and few researchers explore the issue within the field of management (Baba & HakemZadeh, 2012), especially how to build data-driven decision making capability. More research is needed to explore how to help organizations build the data-driven decision making capability (Aksoy, 2013; Goeken, 2011). Conceptualization of data-driven decision making capability may be the foundation to drive such research forward.

It is interesting to note that the concept of DDDM has been thoroughly researched in an educational system setup. Very limited research has been implemented so far though in large organizations to describe a structure of processes needed to be followed in order to determine if the decision makers are as data driven as they want to be.

Brynjolfsson et al. (2011) conclude that DDDM is associated with high productivity. Furthermore, profitability regressions found that DDDM is associated with higher ROE and better asset utilization, but not with increases in ROA or profit margin. Lastly, they found that DDDM can account for a portion of a firm's market value.

Power (2013) states that, today, data-driven decision support is used for a variety of purposes including operational and strategic business intelligence queries, static and real-time performance monitoring, and customer relationship management. He touches upon the issue of DDD Support Systems and concludes that there is a need to identify what decisions will be supported and who might use the proposed data-driven DSS. A powerful sponsor increases the chances that an enterprise-wide DSS will be successfully built and deployed. Further, he claims that it is generally advisable to hire outside expert advice for the first project, referring to an older paper of his (Solomon, 2005; Power, 2002a).

Light et al.(2006) confirm that nowadays we are data rich but information poor. A conceptual framework for data-driven decision making is founded and is used in a notion to depict what it means to be data-driven. With this framework, we shall be able to test the situation within KLM by bringing it down to operational level, contextualizing the levels of data that are being used as an input.

Ikemoto and Marsh (2007) verify the findings of Light et al. (2006) and argue that DDDM varies along two continua: the type of data used and the nature of data analysis and decision making. These types of analyses and decision making also vary along the dimensions of interpretation, reliance on knowledge, type of analysis, extent of participation and frequency.

2.2 DDDM application on for-profit large scale organizations

The ultimate goal of data science is improving decision making, as this generally is of paramount interest to business (Provost & Fawcett, 2013). More organizations are considering how to run smarter, more agile, and more efficient businesses by using the right data to support efficient and effective decision making (Davenport, 2006, Jia et al, 2016). This is generally known as data-driven decision making which emphasizes making decisions based on the analysis of data rather than purely on intuition (Provost & Fawcett, 2013).

Economist Erik Brynjolfsson and his colleagues from MIT and Penn's Wharton School, recently conducted a study of how DDDM affects firm performance. They developed a measure of DDDM that rates firms as to how strongly they use data to make decisions across the company. They show statistically that the more data-driven a firm is, the more productive it is, even controlling for a wide range of possible confounding factors. And the differences are not small: one standard deviation higher on the DDDM scale is associated with a 4–6% increase in productivity. DDDM also is correlated with higher return on assets, return on equity, asset utilization, and market value, and the relationship seems to be causal. DDDM is also associated with significantly higher profitability and market value (Brynjolfsson, Strength in Numbers, 2011).

It is noteworthy to mention that 29% of marketing leaders do not have enough customer data to perform data-driven decision making, and 39% of organizations that collect a large amount of data do not have the capability to convert their customer data into actionable insights (Kumar et al. 2013, Jia et.al 2016). This leads us to understand that DDDM shall be one of the great interests of managers in the upcoming years; however, much past literature on data-driven decision making has not contributed in a practical manner to help organizations build routines of data-driven decision making (Garvin, 2013).

Key findings of selected research on DDDM regarding for-profit organizations via interviews and surveys have shown that:

- Top-performing organizations have used analytics five times more than lower performers. Leading obstacle to widespread adoption is lack of understanding of how to use data to improve the business and a lack of management bandwidth due to competing priorities.
- Data used might vary in the way that is collected, points in time (one time versus longitudinal), type, and level of detail. Analysis and decision making also vary and organizations can be characterized as basic, analysis-focused, data-focused or inquiry-focused.

(Data and Decision Making Same organization, Different perception (Maxwell et.al, 2015)

2.2.1 Perception of DDDM by decision makers and possible benefits

Economic theory suggests that decision makers have to weigh the amounts of information that they receive more than they used to, as the transition from the intuitive management would improve the quality of their decisions. Objective, fine-grained data are now replacing HiPPOs (Highest Paid Person's Opinions) as the basis for decision-making at more and more companies (Kohavi et al., 2009). Furthermore, it is common for companies to purchase a "business intelligence" module to try to make use of the flood of data that they now have on their operations (Brynjolfsson, 2011).

In contrast with the "organizational inertia" that used to prevail in large companies the former decades, having standardized business processes and being reluctant to make important changes nowadays will eventually lead to a disadvantage in the competitive scene of the market. Decision makers themselves sense that they need to acquire more insights through data, for the reason that they may trigger business opportunities that were not evident beforehand.

A more recent study (Lavalle et al., 2010) has reported that organizations using business information and analytics to differentiate themselves within their industry are twice as likely to be top performers as lower performers. Understanding the fundamental concepts, and having frameworks for organizing data-analytic thinking, not only will allow one to interact competently, but will help to envision opportunities for improving data-driven decision making or to see data-oriented competitive threats.

Firms in many traditional industries are exploiting new and existing data resources for competitive advantage. They employ data-science teams to bring advanced technologies to increase revenue and to decrease costs. In addition, many new companies are being developed with data mining as a key strategic component.

Data-science projects require close interaction between the scientists and the business people responsible for the decision making. Firms in which the business people do not understand what the data scientists are doing are at a substantial disadvantage, because they waste time and effort or, worse, because they ultimately make wrong decisions.

LeRoux and Wright (2010) use a broader DDDM framework to examine an organization’s reliance on performance and output indicators, including customer satisfaction and industry standards; however, their survey was not designed to understand how these individual DDDM components combine—or do not combine—to form a systematic process for collecting, analyzing, and using data to make decisions. If stakeholders hold different perceptions about an organization’s DDDM activities and use, research relying on a single individual in that organization might be inaccurate and building a body of knowledge about DDDM—either within an organization or across nonprofits—would require information from multiple individuals in an organization. (Maxwell et al, 2015).

Younger firms tend to show higher rates of productivity growth due to their higher innovation quality (Huergo and Jaumandreu, 2004). Older firms, with well-established routines and business practices and thus a higher degree of organizational inertia, may find the adjustment cost to adopt DDDM too high. Firm age is, therefore, likely to be negatively correlated with DDDM (Brynjolfsson, 2011).

In general, DDDM is associated with higher productivity, higher ROE and better asset utilization, but not with increases in ROA and profit margin; it may also account for a portion of a firm’s market value (Brynjolfsson, 2011).

2.3 DDDM in non-software and/or non-profit related industry

Data-driven decision making (DDDM) has become an emerging field of practice for school leadership and a central focus of education policy and practice (Light, Honey, & Mandinach, 2006). As many educators say, they are data rich, but information poor. By this they mean that there is far too much information with which they must deal, but those data are not easily translatable into information and actionable knowledge. (Light, Honey, & Mandinach, 2006). DDDM in education typically refers to teachers, principals, and administrators systematically collecting and analyzing data to guide a range of decisions to help improve the success of students and schools (Ikemoto and Marsh, 2007).

Profits and revenues measure success in for-profit firms, and these organizations fail when they cannot generate revenues to cover costs. Meanwhile, donations, expenditures, and operating expense ratios have historically been used to capture success in nonprofit organizations (Kaplan, 2001).

Differences in perceptions suggest that building an organization's reliance on DDDM must begin by building a common understanding about what activities are—or are not—being undertaken and that results from research on DDDM using information from only one respondent in an organization might not be reliable (Maxwell et al., 2015).

If data is not collected, organizations cannot analyze information to draw conclusions. If data is not analyzed consistently and correctly, staff might use it to draw incorrect conclusions. Finally, if the results of data analysis are not incorporated fully into decision making, the money spent on collection and analysis is for naught (Maxwell et al., 2015).

DDDM in education specifically refers to teachers, principals, and administrators systematically collecting and analyzing various types of data, including input, process, outcome and satisfaction data, to guide a range of decisions to help improve the success of students and schools (Marsh, 2006). Educators meant very different things when they claimed to be using data or practicing DDDM. (Ikemoto and Marsh, 2007).

Advances in school networking infrastructures and online data warehousing have made it feasible to create systems that use assessment data to support decision making, by providing timely information and presentation and analysis tools to educators across multiple levels of the system (Mandinach et.al, 2006). Several barriers to the effective use of data by educators have been identified (Lim, 2003) including access issues, technical expertise, and training (Choppin, 2002; Cromey, 2000; Mason, 2002; Wayman, 2005).

Further, we know very little about the cognitive strategies teachers employ to transform data into useable information and practice (Herman & Gribbons, 2001). Based on the experience of others (Confrey & Makar, 2005), teachers need to develop fluency in a number of areas in order to make effective use of data. Moreover, many teachers also lack the requisite training to understand, analyze, and connect data to classroom practice. For example, in schools identified as innovative data users, only 19 percent of teachers and school leaders felt they had the requisite knowledge and abilities to manipulate data in meaningful ways (Supovitz & Klein, 2003).

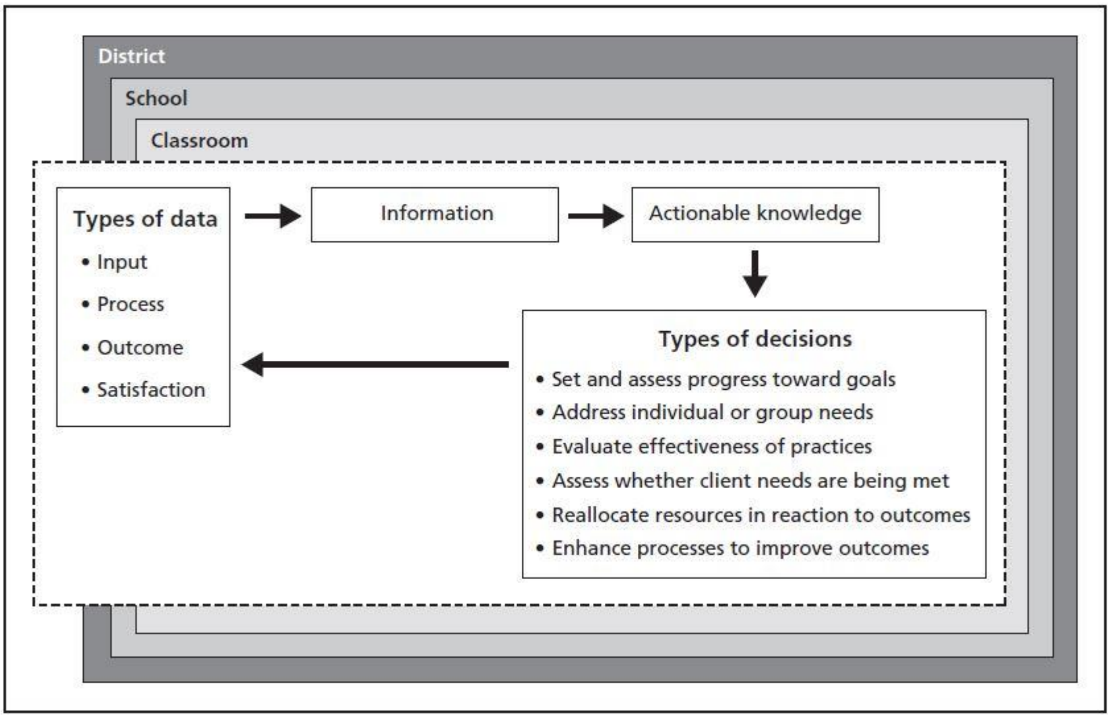

Figure 2: Conceptual framework of Data-Driven Decision Making in Education, Marsh et Al. (2006)

Educators must have specific uses in mind when examining data, and the decisions they make must be both strategic and timely (Mandinach et.al, 2006). While all forms of data use required capacity to translate data into information and actionable knowledge, more complex models of DDDM required additional skills, such as being able to craft good questions, design data-collection instruments (such as surveys), disaggregate and analyze existing data to address new questions, and critique research and other forms of knowledge (Ikemoto and Marsh, 2007).

Teacher anxiety is a contributing factor also to teacher resistance to DDDM and may be an impediment to these teachers engagement in DDDM. (Dunn et al., 2013).

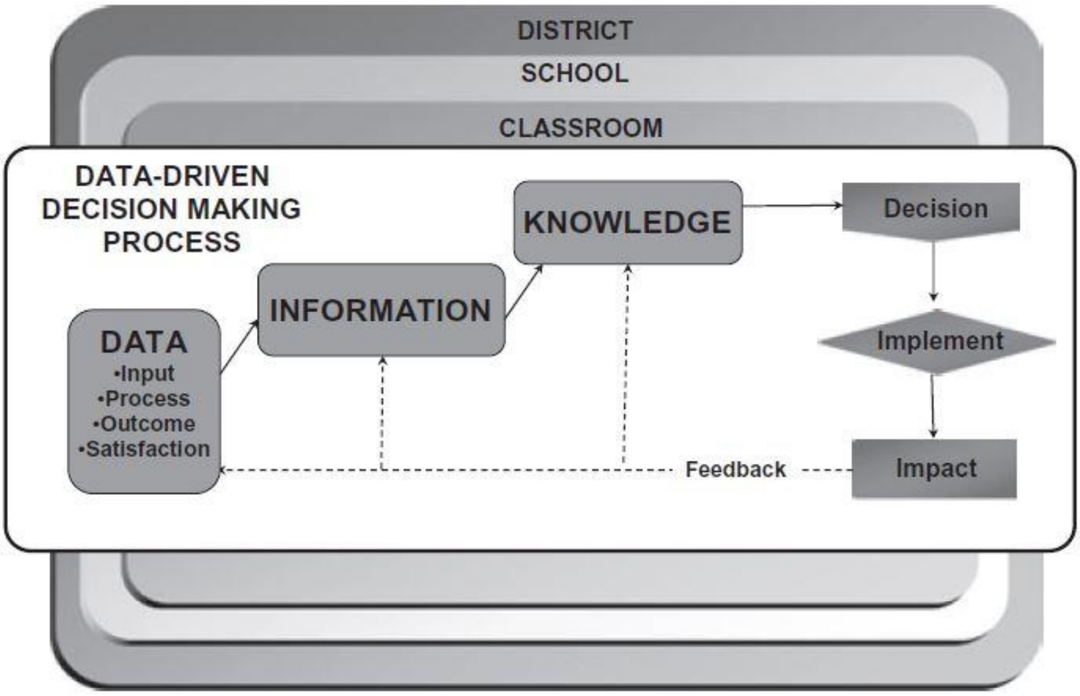

Figure 3: Framework for describing DDDM process in education, Ikemoto and Marsh (2007)

DDDM in practice is not necessarily as linear or continuous as the diagram depicts. For example, educators might skip a step or two in this process by relying on intuition; decide to pause the process to collect additional data; draw on one data source or multiple data sources; or engage in the process alone or as part of a group (Ikemoto and Marsh, 2007).

Given the additional time and capacity required by DDDM, schools and districts were more likely to engage in DDDM—both basic and complex data use and analysis—when external organizations, such as universities, consultants, and state departments of education, were available to help them by providing valuable technical assistance and needed resources (Ikemoto and Marsh, 2007).

That is, the data used in DDDM might vary in the way they were collected, the points in time they represent (one time versus longitudinal), their type (outcome, process, input, satisfaction), and the level of detail and comprehensiveness (aggregated versus disaggregated) (Ikemoto and Marsh, 2007).

Analysis and decision making based on these data can also vary in the way they are conducted (collective versus individual), the extent to which they rely on evidence, expertise, and sophisticated analysis techniques to explain data patterns and identify next steps, and the frequency of the work over time. These decisions generally fall into two categories: decisions that entail using data to inform, identify, or clarify (e.g., identifying goals or needs) and those that entail using data to act (e.g.,changing curriculum, reallocating resources) (Marsh, 2006).

Implications for educators to pursue DDDM would be to acknowledge that DDDM is not a straightforward process, data needs to be improved in a sense of availability, timeliness and comprehensiveness and that educators have to be assisted in order to access external partners, expertise and tools (Ikemoto and Marsh, 2007).

Key findings of selected research on DDDM regarding non-profit organizations via interviews and surveys have shown that:

- There are reported high levels of data collection which do not relate to use.
- Most of the participants stated that they do evaluation using internal resources.
- Programs may provide process data but not client improvement.
- Greater reliance on performance measures increased effectiveness of strategic decision making but client or customer satisfaction and industry standard benchmarks were not so related, suggesting nonprofit managers overlook such information in strategic decision making.
- There was a significant relationship between (1) agencies using performance measurement and the requirement to do so by an outside source; and (2) agencies currently using performance measures and those willing to recommend that others use them (Maxwell et al.,2015).

2.4 DDDM frameworks

DDDM activities may be conceptualized as sequential events in which organizations (1) collect the data needed to make decisions that enhance their services delivery and business operations, (2) analyze the data collected in a manner that they can be verified and used to make decisions, and (3) use data systematically to drive decision making (Maxwell et.al, 2015).

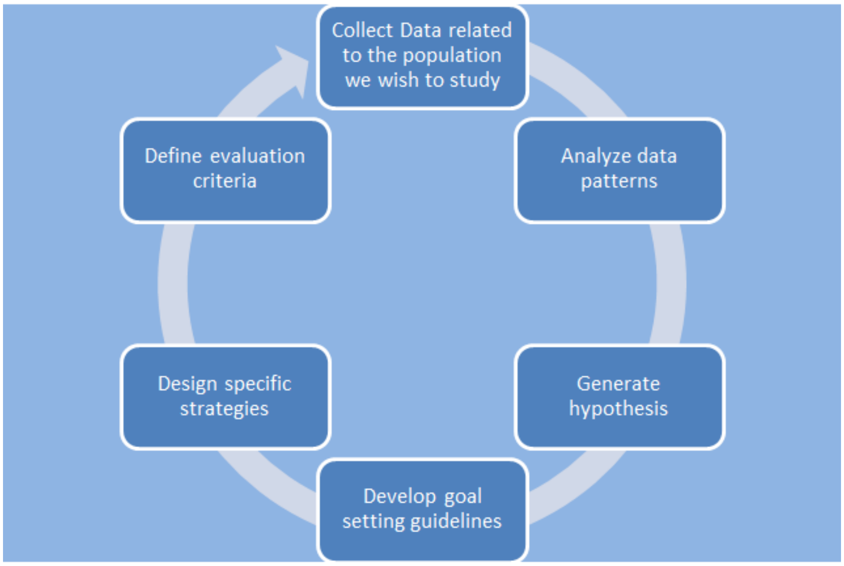

Figure 4: Multiple steps to effectively utilize data to inform decision making, Maxwell et al. (2015)

According to the following framework, data-driven decision making capability is composed of data governance capability, data analytics capability, insight exploitation capability, performance management capability, and integration capability. The process model considers collecting data, processing data into information, transferring data or information into insight, applying insight to decision making, and acting based on performance management. The process also considers the integration of IT infrastructure, process, and people (Hall and Song, 2015).

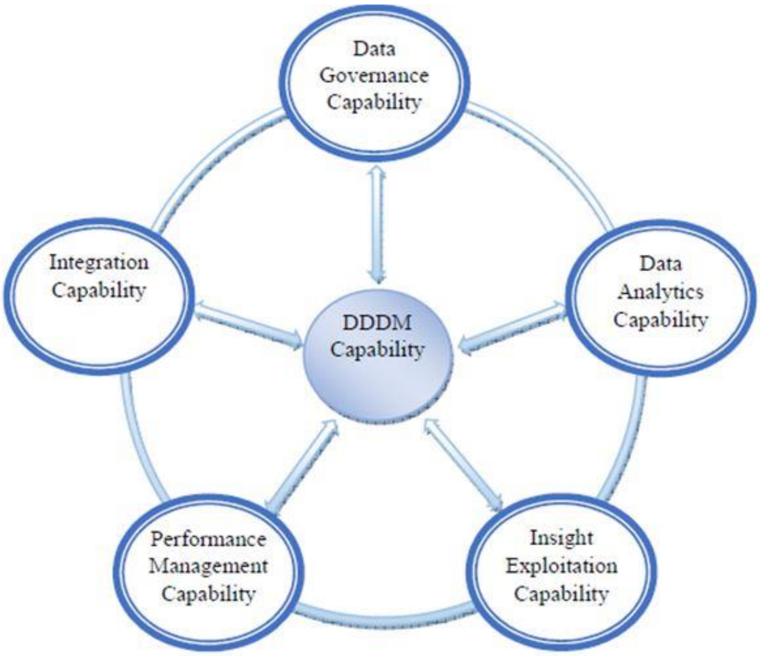

Figure 5: The framework of Data-Driven Decision Making Capability, Hall and Song (2015)

Data driven decision making is especially useful and important when a business owns large datasets that are interconnected and that include time-series data reflecting past, current, and subsequent performance (Morrel-Samuels et al., 2009). Top management should realize the priority importance of data-driven decision making in their operation of business (Aksoy, 2013).

Kumar et al. (2013) reported that 29 percent of marketing leaders do not have enough customer data to perform data-driven decision making, and 39 percent of organizations that collect a large amount of data do not have the capability to convert their customer data into actionable insights. Manyika et al. (2011) also posited that there will be a shortfall of 1.5 million managers with knowledge of performing data-driven decision making by 2018 in the U.S. Thus, it is time to treat data-driven decision making as one responsibility of managers and help them recognize the importance of data-driven decision making and support investment in building their data-driven decision making capability (Aksoy, 2013).

Much past literature on data-driven decision making has not contributed in a practical manner to help organizations build routines of data-driven decision making (Garvin, 2013). More research is needed to explore how to help organizations build the data-driven decision making capability (Aksoy, 2013; Goeken, 2011). Conceptualization of data-driven decision making capability may be the foundation to drive such research forward (Jia, Hall, Song, 2015)

Data-driven decision making capability is defined as the abilities of an organization to utilize data, information, and insight assets in a series of coordinated decision making processes in order to support, inform, or make decisions (Jia, Hall, Song, 2015). The definition summarizes three different roles of data, information, and insight assets in decision making, which are to make a decision, inform a decision, and support a decision (Tingling & Brydon, 2010).

The process model considers collecting data, processing data into information, transferring data or information into insight, applying insight to decision making, and acting based on performance management. The process also considers the integration of IT infrastructure, process, and people. The article of (Jia, Hall, Song, 2015) serves as the foundation of data driven decision making capability research and also encourages more research on how to build an organization's data-driven decision making capability.

According to Ackoff (1989), data, information, and knowledge form a continuum in which data, are transformed to information, and ultimately to knowledge that can be applied to make decisions. As Light and colleagues (2004) note:

- "Data exists in a raw state. It does not have meaning in and of itself, and therefore, can exist in any form, usable or not. Whether or not data becomes information depends on the understanding of the person looking at the data.
- Information is data that is given meaning when connected to a context. It is data used to comprehend and organize our environment, unveiling an understanding of relations between data and context. Alone, however, it does not carry any implications for future action.
- Knowledge is the collection of information deemed useful, and eventually used to guide action. Knowledge is created through a sequential process. In relation to test information, the teacher's ability to see connections between students' scores on different item-skills analysis and classroom instruction, and then act on them, represents knowledge."

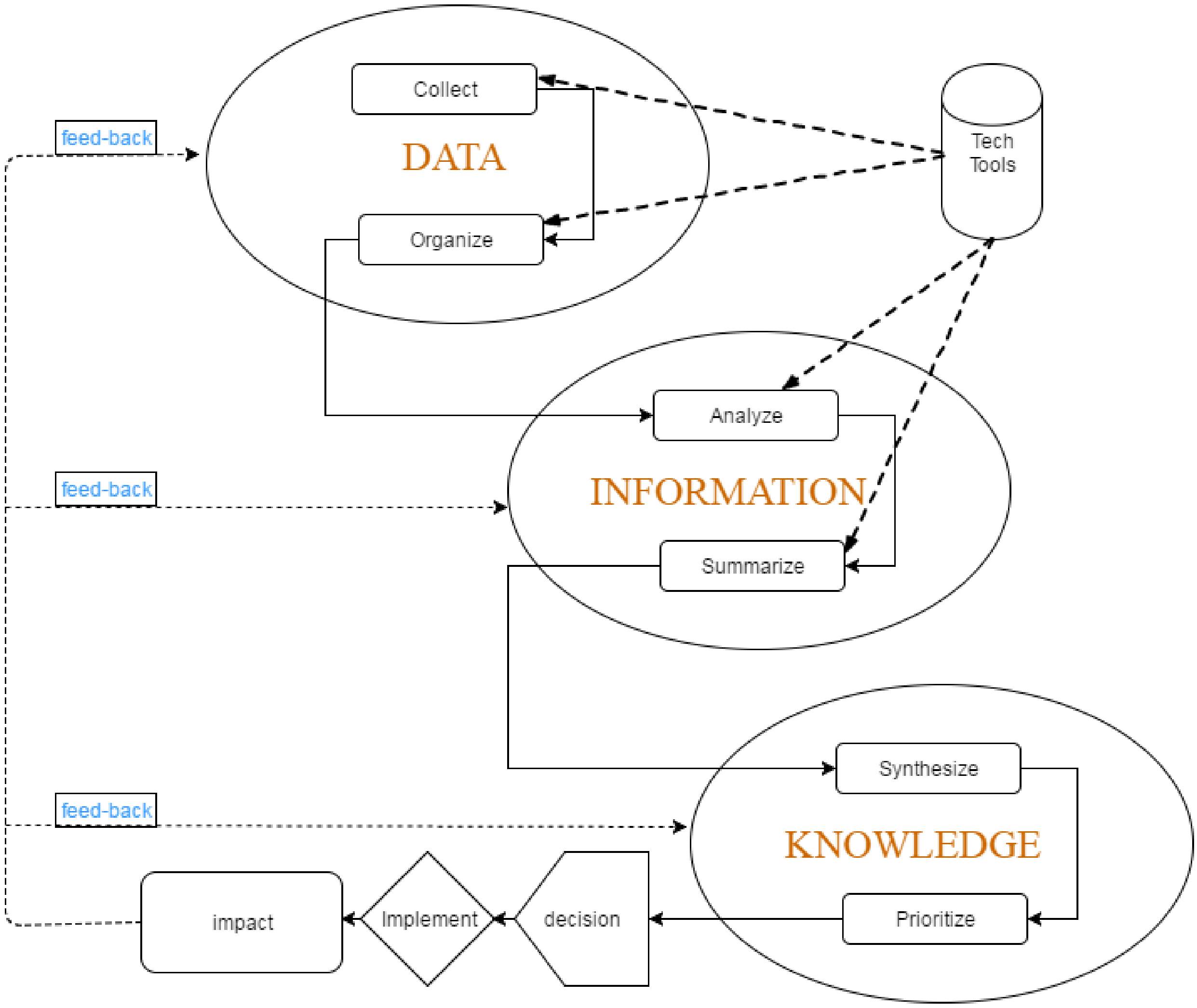

Figure 6: Conceptual model for Data Driven Decision Making

As can be seen in Figure 6, the data to knowledge continuum is defined by the inclusion of six cognitive skills or actions that we have identified as crucial to the decision making process. At the data level, the two relevant skills are “collect” and “organize”. The skills at the information level are “analyze” and “summarize”. At the knowledge level, “synthesize” and “prioritize” are the skills seen as relevant (Mandinach et.al, 2006).

To turn information into knowledge, the stakeholder must synthesize the available information. The final step is to prioritize the knowledge. Setting priorities often requires imparting a value judgment on the accumulated information and knowledge (Mandinach et.al, 2006).

2.5 Case study within KLM - Royal Dutch Airlines

The outcome of this six-step process, moving from data to information to knowledge is a decision. The decision is then implemented, or in some instances may fail to be implemented for other external reasons, such as a lack of resources. The implementation results in some sort of outcome or impact. Depending upon the impact, the decision maker may decide that he or she needs to return to one of the six cognitive steps, thereby creating a feedback loop. The stakeholder may need to collect more data, may need to reanalyze the information, or resynthesize the knowledge. Because of the feedback loops, data-driven decision making is seen as an iterative process with data leading to a decision, implementation of that decision, determination of the impact, and perhaps the need to work through some or all of the six processes again (Mandinach et.al, 2006).

For the purposes of this research, we will examine thoroughly the first level of this three-tier framework that has been introduced; which kind of data is used in the first tier (Data) within the organization under study, KLM - Royal Dutch Airlines. The initial idea was to come across decision makers within the Digital Department of this organization and to collaborate with them in order to figure out in which way can we inform their decisions.

In the aforementioned department, cutting edge technologies are being implemented in order to deliver high-tech solutions regarding all the touchpoints that refer to the booking process of a flight. These touchpoints may be stationary terminals such as desktops, or ordinary portable devices such as smartphones and tablets. New implementations occur daily regarding each device and many software requirements appear frequently in the desk of the decision makers.

The trigger for this research was to assist a certain group of people, working in Agile software development frameworks. All of the interviewees under this study implemented Scrum methodology of working. “A key principle of Scrum is its recognition that during product development, the customers can change their minds about what they want and need (often called requirements volatility), and that unpredicted challenges cannot be easily addressed in a traditional predictive or planned manner (Henry J. and Henry S., 1993). In that sense, the subjects under study were the people behind these teams that led the decisions for the items that need to be delivered, the Product Owners. Most conspicuously, “a Product Owner is a Scrum development role for a person who represents the business or user community and is responsible for working with the user group to determine what features will be in the product release”.

This group of people should actively assist our research regarding the process they follow in order to decide upon which items they need to deliver over others and how do they inform their decisions. Our contribution would be then to identify a set of metrics that are often taken into consideration when a decision needs to be made, to be able to educate them from where they can get this kind of data for these metrics and how could they integrate them to their decision making process. All these aforementioned steps address the first tier of the following framework, which was adopted from the paper of Hall and Song, 2015.

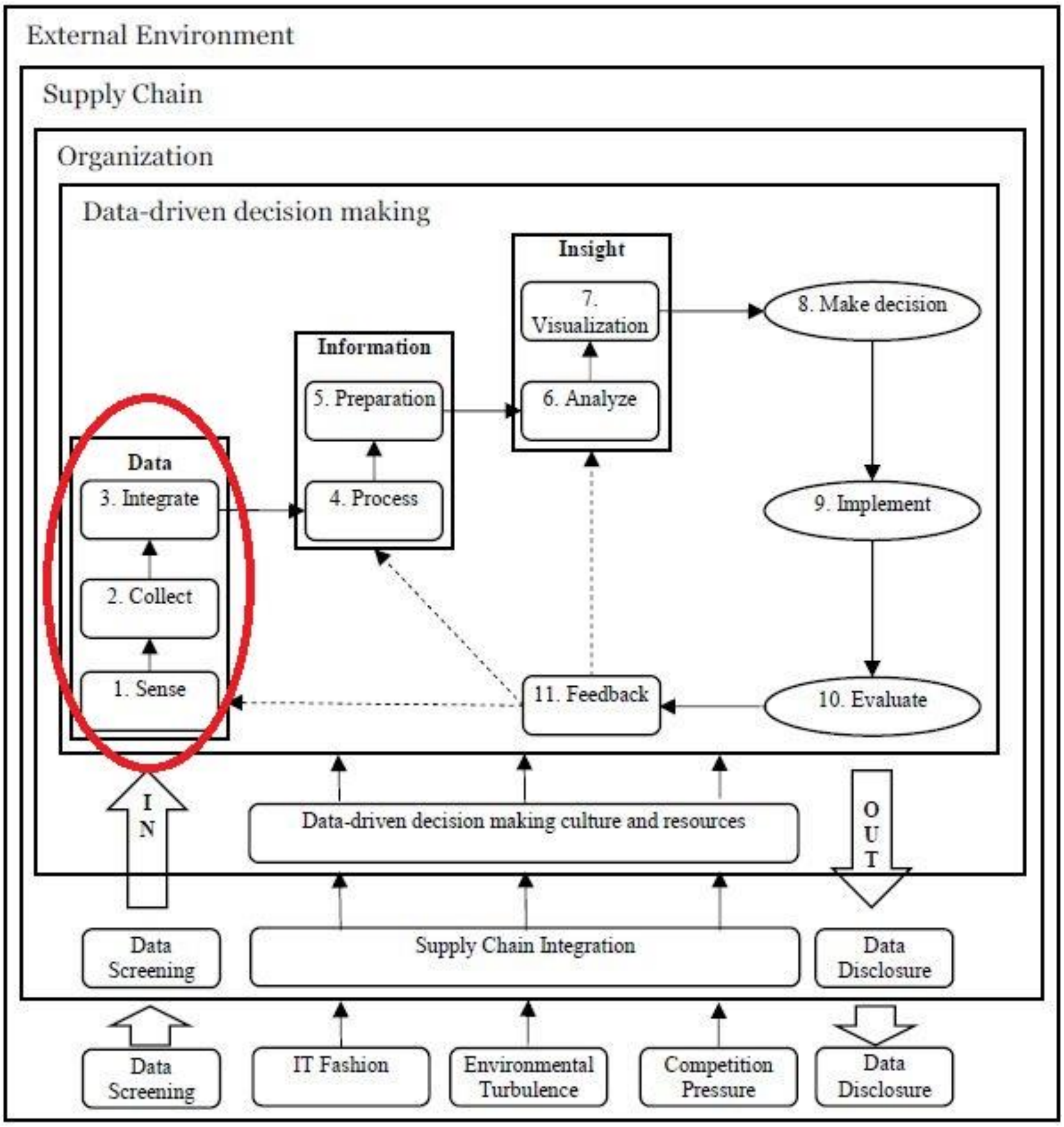

Figure 7: Data driven decision making process, Hall and Song (2015)

2.5.1 Scrum Teams and Product Owners

According to The Scrum Guide by Schwaber and Sutherland (2016), Scrum is a framework for developing and sustaining complex products. Within it, people may address adaptive problems, while productively deliver items of the highest possible value. Since the early 1990's, it has been used to manage complex product development and may employ various processes and techniques. It is founded on empiricism, asserting that knowledge comes from experience and making decisions on what eventually is known. It is mainly characterized by transparency (aspects of the process are available to those responsible for the outcome), by inspection (Scrum users must inspect Scrum artifacts to detect undesirable variances) and by adaptation (adjustments need to be made to minimize deviations).

A Scrum team consists of a Product Owner, the Development team and a Scrum Master. Scrum teams deliver products iteratively and incrementally, maximizing opportunities for feedback. The personas that will be taken under examinations in this current research are the Product Owners (PO). A PO is responsible for maximizing the value of the product and the work of the Development team. He/she is accountable and he/she may represent the desires of a committee in the Product Backlog, but those wanting to change an item's priority within it must address the PO.

2.5.2 Scrum Product Backlog

Schwaber and Sunderland (2016) inform us through 'The Scrum Guide' that a Product Owner is the sole person responsible for managing the Product Backlog (PB). The PB is an ordered list of everything that might be needed in the product and is the single course of requirements for any changes to be made. A PB is never complete; it evolves as the product evolves, giving it a dynamic character. Usually it lists features, functions, requirements, enhancements and fixes that constitute the changes to be made to the product in the future releases. A PB has the attributes of a description, order, estimate and value. Requirements never stop changing, which makes it a living artifact. This is why PB refinement is required; it constitutes the act of adding estimates and order to items and the items may be updated at any time by the PO.

As stated in The Scrum Guide, Product Backlog management includes:

- Clearly expressing Product Backlog items;
- Ordering the items in the Product Backlog to best achieve goals and missions;
- Optimizing the value of the work the Development Team performs;
- Ensuring that the Product Backlog is visible, transparent, and clear to all, and shows what the Scrum Team will work on next; and,
- Ensuring the Development Team understands items in the Product Backlog to the level needed.

2.5.3 Definitions of recognized metrics

Inspired by the Scaled Agile Framework (SAFE) and specifically from the algorithm that is implemented in the Weighted Shortest Job First (WSJF), we recognize that a job's priorities are based on business context, value, time, development facts and risks; automatically ignoring sunk costs, a key principle of Lean economics. In this context, the primary categories that are recognized are user value, business value, time criticality and risk reduction-opportunity enablement value.

Based on these aforementioned categories, our research is oriented into identifying key metrics that may comply within each one. Thus, four categories were initially assumed as top level metrics, within each of those we tried to find out which metrics are mostly taken under consideration by decision makers in the first set of interviews, as described in the third chapter of this study. The algorithm of calculation also inspired the way that the third set of interviews was conducted, scaling metrics according to the Fibonacci sequence, which is also explained further on.

A list of basic definitions shall follow, in order to perceive the nature of this study and to better understand the concept of each recognized metric that was used in order to conduct the interviews.

- *Value* is any desirable result for a stakeholder in any context. It is the ratio of perceived benefits compared to the price for a product or service.
- *Stakeholders* are individuals or groups of people, such as shareholders, customers, employees, suppliers etc., with an interest in the success or failure of a business or project. They can have significant impacts on decisions regarding the operations and finances.
- *Customer* is a party that receives or consumes our goods or services and has the ability to choose between different products and suppliers.

In the context of user value, relevant definitions that need to be addressed are the following:

- *User* is someone who has the authority to use an application, equipment, facility, process or system, or one who consumes a good or service to obtain a benefit or to solve a problem.
- *User satisfaction* refers to the degree to which a product or service meets the user's expectations.
- An *impact* on a user is the assessment of the pros and cons of acting in light of its possible consequences, or the extent and nature of change it may cause.
- *B2e satisfaction* refers to the levels of satisfaction achieved in the exchange of intra-firm information with employees over the internet/intranet.
- *Customer intimacy* is a marketing strategy where a service supplier gets close to the clients. The benefits of greater customer intimacy for a business include improved highly tailored problem solving capabilities and greater adaptation of products to customer needs, as well as higher customer loyalty levels.

In the context of business value, we need to provide the following definitions:

- *Brand equity* is the perceived value of a known name, logo or other identifier. Brand equity affects an organization's ability to market products and services that brand represents.

- *Savings* are the amount of resources not spent on the stage of implementation of a requirement for the final product. It also refers to the portion of disposable income not spent on consumption of goods.
- *Revenue* is an increase in assets or decrease in liabilities caused by the provision of services or products to customers. It is a quantification of the gross activity generated by a business.
- *Partner value* is the sum of benefits that a shareholder acquires from the implementation of the requirement of the final product, not adding value to us directly but to our partners.
- *PR value* refers to the practice of creating and maintaining goodwill of an organization's various publics (customers, employees, investors, etc.), usually through publicity and other non-paid forms of communication.
- *Corporate strategy* is the overall scope and direction of a corporation and the way its various business operations work together to achieve particular goals (alignment with the vision and mission of the company).

In the context of risk reduction, we need to provide the following definitions:

- *Business opportunities* refer to a potentially favorable condition in which a business can capitalize on a changing trend or an increasing demand for a product. It is an ongoing opportunity to generate income as an independent representative of a network marketing company.
- *Security risk* refers to the importance of maintaining the availability of a system, the integrity of data housed by that system and the confidentiality of sensitive information stored on that system. Refers also to the various events that could compromise security of the company and can cause adverse impacts on the organization's business processes or mission.
- *Value of received information* adds to the knowledge of decision makers for future consideration, retrospection and insights.
- *Compliance risk* is exposure to legal penalties, financial forfeiture and material loss an organization faces when it fails to act in accordance with industry laws and regulations, internal policies or prescribed best practices.

Finally, in the context of time criticality, we need to provide the following definitions:

- *Fixed deadline* refers to a date, associated with a plan, that cannot be moved or changed during the schedule.
- *Competitor's initiation* refers to the need of a requirement's implementation that results from the possibility of the scenario that competitors will move towards to this solution before us. Customers might look up for alternative products from other companies.
- *Dependencies to/from others* refer to the degree that this functionality presents dependencies to/from other parties (e.g. teams, features, projects etc.).

3. Research methodology

This chapter aims to present the research methodology and research design that was implemented during this study. In the following sections we analyze the research method that was used, affiliation with the setup of the study and the roles of the participants, the research design that was used in order to achieve all of the quantitative and qualitative data gathering, and eventually the evaluation of this process.

3.1 Conceptualization

Taking into consideration the research 'onion' of Saunders, Lewis and Thornhill (2011), we need to address the different elements of this particular research design. First of all, the procedure followed in order to acquire the relevant information was data collection and data analysis. The time horizon of this research was cross-sectional, as it was implemented within the timeframe of the researcher's internship in the organization of KLM - Royal Dutch Airlines. The strategy that was followed was an embedded case study methodology with the nature of an exploratory study, where we interviewed experts in the subject and conducted in depth individual interviews, as well as with the nature of an explanatory study in the process, where we concluded to several causal relationships between the recognized metrics. At this point, we must mention that the research has the character of embedded mixed methods, where a methodology is embedded within the other during a single means of collecting data; that would be that questions that were in the questionnaires required also qualitative responses.

The latter point can be broken down even further: due to its nature, this study involved exploratory and explanatory research as mentioned. Thus, we observe a triangulation of multiple sources of data, as we used different data collection techniques within the study to ensure that the data tell us what we think they are telling us (Saunders, Lewis and Thornhill, 2011).

During the sets of interviews that were conducted and which will be broken down further in this paper, the researcher tried to avoid threats to reliability. Participant errors were minimized as the interviews were scheduled a fair amount of time beforehand. Furthermore, participant bias was avoided, as the interviews were conducted in a personal room always excluding other participants than the interviewee and the interviewer. Researcher error and researcher bias though may constitute a threat to this study, as all of the interviews had to be transcribed and interpreted by the researcher.

Moreover, validity of the research had to be taken under consideration. At every stage, the researcher tries to avoid mentioning recent events with other interviewees, as well as applying instrumentation, where changes in the layouts of the questionnaires would affect the comparability of results. Mortality could not be avoided, as two of the participants could not conclude the series of interviews until the end, skipping the last part of the series. Finally, ambiguity about causal direction was avoided, as it was clear throughout the whole process which were the causes and the effects of this study. All in all, internal validity was taken under consideration and, in order to be able to apply external validity of the results outside this organization in some certain degree, the results are presented in a coherent way, representing a group of people within a certain department of the organization.

3.2 Research design

The present study follows the process design that is depicted in the figure below.

Figure 8: Process design diagram

At first, the seed of this research topic was planted with discussions of the researcher's manager on how can a team, and in broader perspective an organization, be more data driven. In that sense, the researcher investigated related topics that examined applications of DDDM in fiscal and non-profit organizations. At this stage, in collaboration with several managers of the department, we arranged an initial meeting with stakeholders that were interested in participating in a project where we could explore possibilities of making the decisions of Product Owners more transparent to their stakeholders. This would be the perfect opportunity to initiate a series of interviews with the parties of interest, only to conclude with our findings in the end.

Systematic Literature Review (SLR) showed that too little information was based on how an organization adopts a DDDM scheme, so from that point, the composure of the questionnaires regarding the interviews was conducted. Eventually, and due to time criticality, three different interviews were conducted with each Product Owner, amounting a sum of 28 final and transcribed interviews.

3.2.1 Introductory set of interviews

During the fourth month of the researcher's internship within the organization, the first sets of interviews were concluded. In this stage, a questionnaire with four different sections was used in order to coordinate a semi-structured interview with the Product Owners. In the first section, an introductory conversation was due to be held, in order to familiarize with the interviewee, understand the domain and the amount of responsibilities that they handle.

In the second section, we are trying to get information about the software Requirements Prioritization that occurs within their Product Backlog. With these sets of questions, the researcher aimed to understand which are the main drivers that are behind the decisions that they make, and in which way this procedure occurs.

In the third section, we put under examination the consumers of these products and the stakeholders affiliated with the products. It is often realized that in many cases the consumers were internal teams, which used these deliverable items in order to process their own requirements, creating also dependencies amongst each other. Stakeholders usually have a saying in the delivery process, which also is an interesting topic to investigate and understand to which extent that influences their decisions.

Finally, in the fourth section, we examine if the Product Owners have familiarized themselves with methods or calculations already known, in order to assist their decision making process.

3.2.2 DDDM interviews

During the fifth month of the researcher's internship, all of the sets of interviews regarding Data Driven Decision Making were concluded. The questionnaire was formed based on the feedback of the first set of interviews, and it was enriched after the initial planning. It comprises of four sections also and it examines the literacy of the stakeholders with the topic, obtaining qualitative data, and it requires scoring in different scales from the interviewees regarding their feeling about the usage of the proposed metrics, obtaining quantitative data.

In the first section, the interviewee provides information regarding the types of data that are used within their domain, which steps are followed in order to make business oriented decisions and what are the most significant impediments that they face in order to properly use this kind of data. This process will eventually assist us to perceive how much data-literate are the decisions of Product Owners and in which way can we suggest that they may be enhanced.

In the second section, the interviewee is asked to rate in a numerical scale the level of importance of several recognized metrics that are taken under consideration and may influence a decision. The participants are also required, along with the quantitative data that they input, to explain where would they need to address in order to acquire this kind of data, and in which level are these metrics accessible to them. A brief explanation of these metrics is offered to them, in order to assist them understand what exactly are the purposes of this questionnaire and to reduce the bias or the misinterpretation of the aforementioned drivers.

In the third section, the participants are also asked to provide input in a numerical scale. The purpose of this section is to investigate their technology fluency, their eagerness to examine a more data driven way of thought, their perception about the confidence of their decisions and if they trust that their procedures may be more transparent to their stakeholders.

Lastly, in the fourth section, the participants are required to input numerical answers in a pre-defined scale, inspecting three different sub-sections. Each of those sections refers to a level of the framework of Hall and Song, (2005): collection and organizing (data), analyzing and summarizing (information), synthesizing and prioritizing (knowledge). The purpose of this section is dual; at first, with this way the participant forms an understanding of a data driven framework, and which are the collaborate steps that need to be taken in order to transform raw data into knowledge. Moreover, we are able to understand on which level of this framework the participants are more literate and what kind of inputs they use in order to take decisions. That means, a Product Owner may have to go through hard numbers himself/herself in order to understand what are the benefits of an item's implementation, of if he/she receives knowledge from other parties about the metrics that interest them. Each section contains questions that focus extracting information about the data itself, and in which way it is used by them.

3.2.3 Evaluation of the metrics

During the last round of interviews, all of the participants were required to present a list of their Product Backlogs, containing at least three different items (epics) in order to conduct a prioritization of them with the assist of the recognized metrics. At this stage, all of the Product Owners, except one occasion, prioritized their items based on their intuition, without following any formal way of structuring or informing their decisions.

Each interview lasted approximately 45-60 minutes. This procedure showed us that indeed, it was much more time consuming than any other known methodology; the benefits of this outcome though were that their decision was transparent to their stakeholders, preventing future discussions with their peers in order to justify their prioritization. Stakeholders of different interests were usually involved, which made this prioritization process a matter of political choices several times, in order to please another team that could benefit from it, time wise. As mentioned before, the dependencies amongst development teams was a fact, and the timely delivery of an item could sometimes be crucial for the outcome of a product’s release.

Having a list of several items of their own unique Product Backlog, each Product Owner was asked to score each metric with a numerical input described by the Fibonacci sequence (0,1,2,3,5,8,13,21). The lower the scoring, regarding a specific metric, would indicate the lower level of importance that a metric would have upon the impact of the item’s delivery. Accordingly, the higher a metric would score in the Fibonacci sequence, the more important it would be and the more added value it would have concerning the implementation of that item. In the end, a median average of all the scorings of the metrics would formulate, to inform the Product Owner which items would need to be delivered before others, as the added value of all of these metrics would show which items score higher than others.

At this point, it is crucial to mention that several metrics could be weighed more than others, as they are mostly taken under consideration when a decision needs to be made; thus it wouldn’t be fair for all the metrics to have equivalent value in the Fibonacci enumerations. For the purposes of this study though, we considered that all of these metrics are equally important, and suggestions for future implementations altering the weights of these metrics are mentioned in Chapter 6, referring to possibilities for future research.

Moreover, during the scoring procedure, the POs were requested to score five(5) statements that describe their way of thinking during their estimations using the Fibonacci sequence. The numerical results that we acquire shall assist us get a feeling on how much do the decision makers base their decisions on facts and data and how much on intuition.

3.2.4 Conclusions

The conclusions of this research regarding the main research question, as well as the subsequent questions may be found in Chapter 6, where we depict all of the outcomes regarding the interviews and the implementations of the metrics evaluation.

4. Results

4.1 Initial interviews results

A total of 10 interviews were conducted at this stage with Product Owners. The transcriptions of the interviews come from audio files that were recorded during the interviews. Approximately, each interview lasted 50 minutes, and it had the form of a semi-structured interview. The main outcomes of these interviews are summarized in the following matrixes.

Product Owner	Years of Experience	Number of Projects	Consumers of products	Metrics for prioritization
PO No.1	6 months	3	Flying Blue Programme/ Digital Department	Urgency of delivery/ Technical value/ Time of implementation/ User value/ Business value/ Customer value/ Usability/ Experience of the customer/ Cost reduction/ Enablement of new business opportunities/ Maintainability/ Bringing new technologies
PO No.2	9 months	2	B2C End users	Customer experience/ Velocity of the team/ Time of implementation
PO No.3	36 months	5	B2C End Users	Customer experience/ Business wishes/ Compliancy-regulations-guidelines/ Customer feedback/ Quality of the product/ Customer satisfaction/ Customer value/ Time to market/ Effort/ Cost of delay/ Revenue
PO No.4	36 months	5	B2C End users/ Revenue management/ E-Acquisition team/ E-commerce team	Customer experience/ Improvement of performance and accuracy

Table 1: Results from initial interviews

Product Owner	Years of Experience	Number of Projects	Consumers of products	Metrics for prioritization
PO No.5	12 months	5	Digital Department	Revenue/ Legal requirements/ Brand image/ Time criticality/ Cost of delay
PO No.6	24 months	1	B2C End users/ Mobile team/ Blue Web	Gut feeling/ Time criticality/ Business value
PO No.7	12 months	1	Search WEB/ Counterparts AirFrance	Connecting new customers
PO No.8	15 months	1	Flight Operations	Revenue drivers/ Financial benefits/ Strategic benefits
PO No.9	3 months	1	Applications Users/ B2C End users	Cost of delay/ Customer complaints
PO No.10	42 months	1	B2E (MyWeb Team, Mobile Team, Flight Guides)/ B2C end users	Gut feeling/ Revenue/ Time deadline/ Complexity

Table 2: Results from initial interviews (continued)

From the results depicted above, we realize that the experience of the interviewees ranges from 3 months until 3 ½ years, with involvement in at least 1 project. Most of the times, the consumers of the products that they deliver are the end users, which in our case are the customers of KLM Royal Dutch Airlines. In several occasions though, we notice that the deliverables of their teams aim different development teams within the organization, which immediately affects their stakeholder groups and the dependencies that are created among these teams in order to deliver an item to the end user.

It is mostly noticed that the main driver that leads the decisions of the POs is value in a broader sense. Value, as defined earlier, is a desirable result for stakeholders in any context; in that sense, the responses that we received could be categorized under the spectrum of user value, hence the customer, which relates to customer experience and customer satisfaction for instance. In another sense, value can be seen from the context of business value, and within we can realize metrics such as revenue, brand value and partner value.

At this stage, we are able to comprehend the main drivers that lead decision makers to conclude to a decision, and with which we continue the series of interviews further on.

4.2 DDDM interviews results

A total of 10 interviews were conducted with the same personas during this stage as well. The interviews were conducted in a more structured way this time, as the interviewer asked the participants to elaborate in topics that were categorized in four different sections. In the first section, we acquire qualitative data from the responses of the POs regarding their literacy concerning their perceptions about the terms and content of DDDM. The interviewees elaborated on the types of data that they go through, how do they use them to arrive to a business related decision, and what would the most significant barriers be, regarding the effective utilization of this data.

In the second section, we switch to a quantitative measurement, asking the participants to enumerate in scales of importance several metrics, that were comprehended during the initial interviews and from ones that came up from literature review. All of the participants provided numbers related to their perception of importance for each metric, and what kind of role does this metric have during their decision making process. It is interesting to note that the POs would address issues on how to find this kind of data for the aforementioned metrics, and if the access to this kind of data is feasible or not. In the following picture, we can see all of the answers acquired from the POs summarized in a sheet.

At this point it needs to be stressed out that the levels of importance are in a scale from 1 to 10, with 1 being of utmost importance to 10 being the of least importance metric.

SECTION 2	P.O. No1	P.O. No2	P.O. No3	P.O. No4	P.O. No5	P.O. No6	P.O. No7	P.O. No8	P.O. No9	P.O. No10	Totals
Levels of importance regarding the delivery of a requirement (Scale 1-10, 1 being most important and 10 being least important)											
Number of impacted customers	3	2	1	1	2	4	2	4	2	2	23
User satisfaction effect	4	4	3	3	2	1	3	2	2	2	26
PR Value	4	4	7	4	7	4	5	4	5	3	47
Brand Value	3	2	7	5	7	3	4	5	5	3	44
Cost Savings	2	1	2	1	3	2	4	3	2	2	22
Revenue	2	1	1	1	3	2	1	3	2	1	17
Partner Value	4	4	8	10	6	4	3	5	8	4	56
New Business Opportunities	3	3	7	9	3	3	2	7	3	2	42
Security Risks	2	1	3	5	4	2	3	2	8	1	31
Potential penalties/Negative effects	4	5	2	6	4	1	5	3	3	3	36
Value of received information	6	2	2	5	5	4	3	4	5	2	38
Fixed deadline	2	1	5	1	3	2	6	2	2	3	27
Competitors Initiation	6	3	5	6	5	10	4	4	3	4	50
Dependencies to/from others	4	2	2	4	3	2	1	2	3	4	27
Business value reduction over time	3	4	4	4	3	(x)	5	4	7	3	37
Corporate Strategy	3	1	2	4	4	3	3	6	7	1	34
B2e Satisfaction	9	10	1	1	8	6	6	4	7	5	57
Loyalty - Repurchase	(x)	4	(x)	1	(x)	(x)	2	3	(x)	2	12
Customer Intimacy	(x)	5	(x)	1	(x)	(x)	2	3	(x)	1	12

Table 3: Levels of importance regarding given metrics

In the third section, we come back to several generic questions regarding their prioritization procedure, only now they enumerate in a scale of importance the statements that are provided to them. During the analysis further on, it is interesting to see their perception regarding their confidence of prioritizing, the transparency to their stakeholders and their belief if these procedures may be enhanced by obtaining more data. A summary of their answers is depicted below.

Here, we need to stress out that the levels of importance are in a scale from 1 to 5, with 1 standing for full agreement and 5 standing for full disagreement.

SECTION 3	P.O. No1	P.O. No2	P.O. No3	P.O. No4	P.O. No5	P.O. No6	P.O. No7	P.O. No8	P.O. No9	P.O. No10		Totals
Quantifiable measurements												
(1 being the highest, 5 the lowest)												
Personal technical fluency	2	2	2	4	1	3	1	2	3	3		23
Confidence for current prioritization procedure	3	2	5	4	3	2	2	2	2	2		27
Willingness to apply new methods to standarize prioritization procedure	1	1	1	1	1	3	4	3	2	2		19
Satisfaction of prioritization transparency to stakeholders	1	4	5	5	2	3	3	3	4	3		33
Belief that transparency to stakeholders can be enhanced	2	2	1	1	1	3	2	1	2	2		17
Understanding of DDDM principles and practices	3	3	2	4	3	2	1	2	3	1		24
Belief that data driven practices may improve business outcomes	2	1	1	2	1	1	1	1	2	1		13

Table 4: Scoring of given statements for quantitative analysis

Lastly, in the final section, the participants provide enumerated answers in a scale of importance regarding the levels of the DDDM framework of Hall and Song (2015), where there are three different levels of the DDDM process; data, information and knowledge. The interviewer collected these responses from the participants in order to get a feeling about each level. Similarly, we need to stress out that the levels of importance are in a scale from 1 to 5. '1' stands for 'Strongly Agree', '2' stands for 'Agree', '3' stands for 'Neither Agree nor Disagree', '4' stands for 'Disagree', '5' stands for 'Strongly Disagree', while now they have the option to mark a question with a 'D' which would stand for 'Do not know':

- For the collection of data, POs answered statements regarding the collection procedures of data and the amount of data available in order to cover all of their customer's segmentations. The questions focused to understand if they, as the spokespersons of their department, archive data properly, have access to their preferred datasets easily and if they are dependent for important customer's data from other departments. All of the answers regarding the data tier are summarized below.

SECTION 4	P.O. No1	P.O. No2	P.O. No3	P.O. No4	P.O. No5	P.O. No6	P.O. No7	P.O. No8	P.O. No9	P.O. No10		Totals
Data tier												
As a department, we have an efficient data collection system in place	2	4	3	5	5	3	2	3	1	4		32
As a department, we have sufficient resources to collect data	4	D	2	4	5	3	3	2	1	3		27
As a department, we collect data for all of our customer's segments	4	D	4	1	5	2	4	D	1	2		23
Data collection methodologies are being conducted properly	2	2	D	1	4	D	3	D	D	3		15
Collecting data is time consuming	1	D	2	1	5	2	1	2	3	5		22
As a department, we have open discussions about data.	4	3	3	3	3	4	1	4	1	1		27
As a department, we archive all the data that we receive properly	D	2	D	5	5	D	4	4	2	D		22
As a department, we have access to preferred datasets easily	D	5	D	4	2	4	2	4	3	4		28
As a department, we are dependant for important-	2	5	2	1	2	2	2	2	4	3		25
-customer's data from other departments												

Table 5: Scoring of given statements regarding data for quantitative analysis

- For the analysis of data, POs answered statements that related to customer service and what kind of added value does the translation of data into actions offers. Their answers may be realized in the following picture.

SECTION 4	P.O.No1	P.O.No2	P.O.No3	P.O.No4	P.O.No5	P.O.No6	P.O.No7	P.O.No8	P.O.No9	P.O.No10	Totals
Infomation tier											
Analysis improves services we provide to customers	1	1	1	1	1	2	1	1	2	1	12
Analysis improves the skills and job performance of employees	1	1	1	1	3	2	2	2	2	2	17
As a department, we have staff with expertise in data analysis	3	1	2	5	4	4	4	2	1	2	28
I have received training to effectively interpret and act upon results	4	1	5	5	4	4	5	3	4	3	38
As a department, we translate discussions of data into actions	4	5	2	5	3	5	3	4	2	2	35
Translating data takes away from the time spent collecting data	5	D	D	D	3	2	2	D	4	D	16

Table 6: Scoring of given statements regarding information for quantitative analysis

- For the prioritization of data, POs scored several statements that related to the efficient usage of data in order to reflect on their decisions and to conceptualize their main focus when they obtain information regarding several metrics. Their replies may be seen in the below picture.

SECTION 4	P.O.No1	P.O.No2	P.O.No3	P.O.No4	P.O.No5	P.O.No6	P.O.No7	P.O.No8	P.O.No9	P.O.No10	Totals
Knowledge tier											
Using data builds an understanding of how our domain operates	4	1	2	2	2	3	4	2	1	1	22
We acquire sets of information from multiple sources	4	1	1	5	1	2	1	2	1	2	20
As a department, we use data to improve the quality of decision-making	3	2	2	5	4	4	3	3	3	2	31
As a department, we use data to reflect on our decisions	4	3	2	5	4	5	5	3	2	3	36
Our main focus is to increase customer satisfaction and quality of service	3	1	1	1	1	4	1	2	2	2	18
Our main focus is to increase the business value constantly	2	1	2	1	2	3	3	2	1	3	20
Our main focus is to reduce potential risks for our domain and organization	3	1	3	4	2	4	4	2	3	D	26
Our main focus is to implement our features and deliver on time	2	1	1	1	1	1	3	2	1	2	15

Table 7: : Scoring of given statements regarding knowledge for quantitative analysis

4.3 PO Matrix interviews results

Following the DDDM interviews, 8 out of the 10 POs volunteered to participate in the prioritization methodology regarding their Product Backlogs, as they were formed in that present timeframe. These lists contained epics that had to be arranged and that they were already in a priority list, according to the PO's perception. Our common goal was to re-prioritize them based on the average score that each epic would have, following the Fibonacci sequence scoring for each metric. The highest average scoring, taking consideration all of the metrics, stands for a software requirement that needs to be prioritized first.

This interview lasted approximately 1 hour with each Product Owner, and while they were scoring for each metric, they were requested to evaluate several statements concurrently. These statements were aiming to inform us about the availability of related data and information about the metric being scored, and if they scored for the specific metric based on their knowledge or based on their intuition. The results of the 8 interviews conducted with the POs are depicted in the following pictures.

The sets of metrics taken under consideration are broken down into four categories. The nomination of these categories is inspired from the Weighted Shortest Job First (WSJB) framework. The concept here is to break down each of these categories in sets of metrics in order to give the opportunity to the decision maker to score in more specific items.

User Value	Business Value	Time Criticality	Risk Reduction
Number of impacted customers	PR value	Business value decrease over time	Potential penalties & negative effects
Effect on user satisfaction	Brand value	Fixed deadline	Security or privacy risk
Loyalty/ Repurchase Tension	Cost savings	Competitor's initiation	Value of received information
	Additional Revenue	Dependencies on/from others	
	Contribution to Corporate Strategy	Enables business opportunities	
	Partner value		

Figure 9: Categorized sets of metrics

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty - Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	Partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies on/from others	Enables business opportunities	Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
User Value																			
Business value																			
Time criticality																			
Risk reduction																			
Single account	5,00	3,00	1,00	1,00	1,00	2,00	1,00	3,00	1,00	1,00	1,00	1,00	1,00	1,00	3,00	1,00	1,00	28,00	1,65
I have based my scoring on my knowledge (experience).	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
My choice is a gut-feeling.	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
I have based my scoring using data.	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00		2,00
I trust that there is related information for this driver.	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00		3,00
I can have access to this information.	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00		2,00
FB ultimate	3,00	3,00	3,00	1,00	3,00	1,00	1,00	1,00	1,00	1,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	29,00	1,71
I have based my scoring on my knowledge (experience).	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
My choice is mainly a gut-feeling.	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00		3,00
I have based my scoring using data.	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
I trust that there is related information for this driver.	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
I can have access to this information.	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00		3,00
FB mini enrolment	1,00	2,00	1,00	1,00	1,00	1,00	1,00	4,00	4,00	1,00	1,00	1,00	2,00	1,00	2,00	2,00	2,00	28,00	1,65
I have based my scoring on my knowledge (experience).	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
My choice is mainly a gut-feeling.	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00		2,00
I have based my scoring using data.	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
I trust that there is related information for this driver.	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
I can have access to this information.	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00		2,00
Mileage summary	3,00	5,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	3,00	1,00	1,00	1,00	1,00	1,00	1,00	29,00	1,71
I have based my scoring on my knowledge (experience).	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
My choice is mainly a gut-feeling.	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00		4,00
I have based my scoring using data.	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00		2,00
I trust that there is related information for this driver.	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00		3,00
I can have access to this information.	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00		2,00

Figure 10: Implementation of framework in the Product Backlog of Product Owner No.1

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty- Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	External Partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies on/from others	Enables business opportunities	IT Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
UserValue																			
Business value																			
Time criticality																			
Risk reduction																			
Branded fares (depiction of different prices to the customer	**21,00**	**8,00**	**8,00**	**8,00**	**5,00**	**1,00**	**21,00**	**3,00**	**0,00**	**1,00**	**0,00**	**5,00**	**21,00**	**3,00**	**0,00**	**1,00**	**3,00**	**109,00**	**6,41**
I have based my scoring on my knowledge (experience).	1,00	3,00	4,00	2,00	3,00	1,00	1,00	4,00	1,00	4,00	1,00	4,00	1,00	5,00	1,00	1,00	3,00		2,35
My choice is a gut-feeling.	5,00	2,00	1,00	2,00	2,00	5,00	5,00	2,00	5,00	1,00	5,00	1,00	5,00	1,00	5,00	5,00	4,00		3,29
I have based my scoring using data.	1,00	5,00	5,00	4,00	5,00	1,00	1,00	5,00	5,00	5,00	2,00	5,00	1,00	5,00	1,00	2,00	4,00		3,35
I trust that there is related information for this driver.	1,00	5,00	0,00	5,00	0,00	1,00	1,00	0,00	5,00	5,00	1,00	0,00	1,00	3,00	1,00	1,00	0,00		2,50
I can have access to this information.	1,00	5,00	0,00	5,00	0,00	1,00	2,00	1,00	5,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		2,13
Group fares (offering a fare to a group)	**3,00**	**5,00**	**5,00**	**13,00**	**1,00**	**2,00**	**2,00**	**1,00**	**1,00**	**2,00**	**0,00**	**8,00**	**1,00**	**1,00**	**0,00**	**0,00**	**1,00**	**46,00**	**2,71**
I have based my scoring on my knowledge (experience).	1,00	4,00	5,00	2,00	5,00	1,00	1,00	1,00	2,00	1,00	1,00	4,00	1,00	1,00	1,00	1,00	1,00		1,94
My choice is mainly a gut-feeling.	5,00	1,00	1,00	4,00	1,00	5,00	5,00	5,00	4,00	4,00	5,00	1,00	5,00	5,00	5,00	5,00	5,00		3,88
I have based my scoring using data.	1,00	5,00	5,00	2,00	5,00	2,00	1,00	5,00	5,00	5,00	3,00	5,00	5,00	1,00	5,00	5,00	5,00		3,82
I trust that there is related information for this driver.	1,00	4,00	0,00	1,00	0,00	1,00	1,00	5,00	5,00	0,00	4,00	0,00	1,00	1,00	1,00	1,00	1,00		2,08
I can have access to this information.	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00		1,06

Figure 11: Implementation of framework in the Product Backlog of Product Owner No.2

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty- Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	Partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies on/from others	Enables business opportunities	Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
UserValue																			
Business value																			
Time criticality																			
Risk reduction																			
As AF/KLM we want to help our passengers who experience any form of disruption, with the right compensation which they can receive via their desired channel of issuance and can be easily re																			
	5,00	**13,00**	**8,00**	**5,00**	**3,00**	**8,00**	**1,00**	**3,00**	**8,00**	**2,00**	**8,00**	**8,00**	**13,00**	**5,00**	**5,00**	**1,00**	**8,00**	104,00	**6,12**
I have based my scoring on my knowledge (experience).	2,00	2,00	4,00	4,00	5,00	2,00	4,00	5,00	4,00	4,00	1,00	4,00	1,00	4,00	5,00	4,00	1,00		3,29
My choice is a gut-feeling.	3,00	3,00	1,00	2,00	1,00	3,00	2,00	1,00	2,00	1,00	5,00	1,00	5,00	2,00	1,00	2,00	5,00		2,35
I have based my scoring using data.	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00		5,00
I trust that there is related information for this driver.	1,00	2,00	3,00	3,00	4,00	1,00	2,00	4,00	2,00	0,00	1,00	1,00	1,00	0,00	0,00	4,00	2,00		2,21
I can have access to this information.	1,00	3,00	3,00	3,00	4,00	1,00	3,00	4,00	2,00	0,00	1,00	4,00	2,00	0,00	0,00	4,00	2,00		2,64
As KLM we want to enable flying blue elite members to buy access to the KLM lounge as a gift with as a result that the reciever can access the lounge without travelling with the FB elite membe																			
	3,00	**5,00**	**5,00**	**5,00**	**5,00**	**0,00**	**5,00**	**3,00**	**8,00**	**1,00**	**21,00**	**2,00**	**8,00**	**5,00**	**1,00**	**0,00**	**3,00**	80,00	**4,71**
I have based my scoring on my knowledge (experience).	4,00	4,00	5,00	4,00	5,00	4,00	4,00	5,00	4,00	5,00	1,00	4,00	3,00	5,00	5,00	4,00	5,00		4,18
My choice is mainly a gut-feeling.	2,00	2,00	1,00	2,00	1,00	2,00	2,00	1,00	2,00	1,00	5,00	2,00	3,00	1,00	1,00	2,00	1,00		1,82
I have based my scoring using data.	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	1,00	5,00	5,00	5,00	5,00	5,00	5,00		4,76
I trust that there is related information for this driver.	1,00	0,00	3,00	0,00	0,00	2,00	2,00	0,00	3,00	0,00	1,00	0,00	2,00	0,00	3,00	4,00	0,00		2,33
I can have access to this information.	1,00	0,00	3,00	0,00	0,00	2,00	2,00	0,00	2,00	0,00	1,00	0,00	2,00	0,00	3,00	4,00	0,00		2,22
As KLM we want to ensure that we show accurate flight statuses to our passengers with as a result that they are reassured about the status of their flight and can take action accordingly.																			
	21,00	**13,00**	**3,00**	**1,00**	**1,00**	**3,00**	**3,00**	**5,00**	**13,00**	**8,00**	**0,00**	**1,00**	**1,00**	**13,00**	**5,00**	**3,00**	**13,00**	107,00	**6,29**
I have based my scoring on my knowledge (experience).	1,00	1,00	3,00	4,00	4,00	4,00	5,00	4,00	2,00	3,00	1,00	5,00	3,00	3,00	2,00	5,00	1,00		3,00
My choice is mainly a gut-feeling.	5,00	5,00	3,00	2,00	2,00	2,00	1,00	2,00	4,00	3,00	5,00	1,00	3,00	3,00	4,00	1,00	5,00		3,00
I have based my scoring using data.	3,00	2,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	2,00	1,00	5,00	5,00	3,00	3,00	5,00	2,00		3,88
I trust that there is related information for this driver.	1,00	1,00	0,00	0,00	0,00	2,00	2,00	0,00	2,00	1,00	1,00	2,00	2,00	1,00	1,00	1,00	1,00		1,38
I can have access to this information.	1,00	1,00	0,00	0,00	0,00	2,00	2,00	0,00	2,00	1,00	1,00	2,00	2,00	1,00	1,00	5,00	1,00		1,69

Figure 12: Implementation of framework in the Product Backlog of Product Owner No.3

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty- Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	Partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies of/from others	Enables business opportunities	IT Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
User Value																			
Business value																			
Time criticality																			
Risk reduction																			
Flight Summary	0,00	0,00	0,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	3,00	0,00	5,00	0,00	1,00	1,00	0,00	12,00	0,71
I have based my scoring on my knowledge (experience).	4,00	4,00	4,00	4,00	4,00	2,00	1,00	3,00	1,00	2,00	3,00	2,00	1,00	1,00	3,00	3,00	5,00		2,76
My choice is a gut-feeling.	1,00	1,00	1,00	1,00	1,00	4,00	1,00	1,00	1,00	1,00	3,00	1,00	1,00	1,00	1,00	1,00	1,00		1,29
I have based my scoring using data.	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00		5,00
I trust that there is related information for this driver.	5,00	5,00	5,00	5,00	5,00	5,00	0,00	5,00	0,00	0,00	2,00	0,00	1,00	0,00	1,00	2,00	0,00		3,73
I can have access to this information.	5,00	5,00	5,00	5,00	5,00	5,00	0,00	5,00	0,00	0,00	2,00	0,00	1,00	0,00	1,00	2,00	0,00		3,73
Rate the Flight	5,00	5,00	5,00	2,00	2,00	0,00	0,00	13,00	0,00	2,00	2,00	2,00	13,00	0,00	1,00	0,00	5,00	57,00	3,35
I have based my scoring on my knowledge (experience).	3,00	2,00	3,00	4,00	4,00	2,00	2,00	2,00	2,00	4,00	2,00	4,00	1,00	3,00	5,00	4,00	3,00		2,94
My choice is mainly a gut-feeling.	2,00	2,00	2,00	1,00	1,00	1,00	1,00	3,00	1,00	1,00	1,00	1,00	5,00	1,00	1,00	1,00	1,00		1,53
I have based my scoring using data.	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00		5,00
I trust that there is related information for this driver.	3,00	2,00	3,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	5,00	1,00	0,00	1,00	0,00	0,00		2,57
I can have access to this information.	3,00	2,00	3,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00		1,83
Check In for MyDuty	0,00	0,00	0,00	0,00	0,00	13,00	0,00	8,00	0,00	0,00	13,00	0,00	13,00	0,00	13,00	0,00	0,00	60,00	3,53
I have based my scoring on my knowledge (experience).	1,00	1,00	1,00	1,00	1,00	1,00	1,00	3,00	1,00	3,00	1,00	1,00	1,00	3,00	1,00	3,00	4,00		1,65
My choice is mainly a gut-feeling.	1,00	1,00	1,00	1,00	1,00	5,00	1,00	3,00	1,00	1,00	5,00	1,00	5,00	1,00	5,00	1,00	1,00		2,06
I have based my scoring using data.	5,00	5,00	5,00	5,00	5,00	1,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00		4,76
I trust that there is related information for this driver.	0,00	0,00	0,00	0,00	0,00	1,00	0,00	3,00	0,00	0,00	2,00	0,00	1,00	0,00	1,00	2,00	0,00		1,67
I can have access to this information.	0,00	0,00	0,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	3,00	0,00	1,00	0,00	1,00	3,00	0,00		1,67

Figure 13: Implementation of framework in the Product Backlog of Product Owner No.4

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty- Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	Partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependend on by others	Enables business opportunities	Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
User Value																			
Business value																			
Time criticality																			
Risk reduction																			
Branded Fares	21,00	5,00	8,00	5,00	5,00	0,00	21,00	21,00	3,00	13,00	13,00	8,00	13,00	21,00	0,00	2,00	8,00	167,00	9,82
I have based my scoring on my knowledge (experience).	2	2	4	5	5	5	3	3	5	3	2	4	2	2	2	3	4		3,29
My choice is a gut-feeling.	3	1	1	1	1	1	1	1	1	3	5	1	3	1	2	1	1		1,65
I have based my scoring using data.	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5		5,00
I trust that there is related information for this driver.	3	5	5	5	5	4	3	2	5	4	4	4	3	3	4	3	4		3,88
I can have access to this information.	2	5	5	5	5	4	4	4	5	5	5	4	3	4	4	3	4		4,18
Native Checkouts	13,00	13,00	5,00	8,00	5,00	0,00	5,00	2,00	2,00	3,00	5,00	8,00	21,00	8,00	8,00	5,00	5,00	116,00	6,82
I have based my scoring on my knowledge (experience).	3	3	4	4	4	3	4	5	5	4	4	4	1	4	4	4	4		3,76
My choice is mainly a gut-feeling.	2	2	1	1	1	2	1	1	1	1	1	2	4	1	1	1	1		1,41
I have based my scoring using data.	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5		4,94
I trust that there is related information for this driver.	3	5	5	4	4	5	5	5	5	5	4	4	3	4	3	4	5		4,29
I can have access to this information.	2	5	5	4	4	5	5	5	5	5	4	3	2	5	3	3	5		4,12
Redesign the press App	5,00	8,00	5,00	8,00	13,00	8,00	2,00	8,00	3,00	5,00	3,00	1,00	13,00	13,00	3,00	3,00	5,00	106,00	6,24
I have based my scoring on my knowledge (experience).	4	4	5	4	4	3	3	4	4	4	3	3	1	2	3	4	4		3,47
My choice is mainly a gut-feeling.	2	2	1	2	2	3	3	3	2	1	3	3	5	4	3	2	2		2,53
I have based my scoring using data.	5	5	5	5	5	4	5	4	5	5	5	5	5	5	5	5	5		4,88
I trust that there is related information for this driver.	3	4	4	4	3	2	3	3	5	4	2	5	1	3	2	5	4		3,35
I can have access to this information.	2	3	4	3	3	3	3	3	5	4	2	4	1	2	3	4	4		3,12

Figure 14: Implementation of framework in the Product Backlog of Product Owner No.5

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty - Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	Partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies on/from others	Enables business opportunities	Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Avg.
UserValue																			
Business value																			
Time criticality																			
Risk reduction																			
Check-In (e-conv site)	21,00	13,00	3,00	2,00	2,00	5,00	1,00	1,00	8,00	1,00	0,00	5,00	8,00	5,00	3,00	21,00	5,00	104,00	6,12
I have based my scoring on my knowledge (experience).	3,00	2,00	2,00	2,00	2,00	3,00	1,00	1,00	2,00	2,00	1,00	2,00	2,00	2,00	3,00	1,00	3,00		2,00
My choice is a gut-feeling.	5,00	4,00	2,00	1,00	1,00	3,00	5,00	5,00	4,00	4,00	5,00	2,00	4,00	4,00	2,00	4,00	3,00		3,41
I have based my scoring using data.	1,00	3,00	3,00	5,00	5,00	3,00	1,00	1,00	2,00	4,00	5,00	5,00	2,00	3,00	3,00	2,00	5,00		3,12
I trust that there is related information for this driver.	1,00	2,00	3,00	5,00	5,00	2,00	1,00	1,00	2,00	4,00	1,00	5,00	2,00	3,00	3,00	2,00	4,00		2,71
I can have access to this information.	1,00	1,00	2,00	5,00	5,00	1,00	1,00	1,00	1,00	5,00	1,00	5,00	1,00	1,00	1,00	1,00	1,00		2,00
Ancillary sales	13,00	8,00	8,00	8,00	8,00	1,00	21,00	21,00	13,00	3,00	0,00	5,00	13,00	21,00	3,00	13,00	8,00	167,00	9,82
I have based my scoring on my knowledge (experience).	3,00	2,00	3,00	3,00	3,00	1,00	1,00	2,00	3,00	3,00	1,00	2,00	2,00	1,00	2,00	2,00	3,00		2,18
My choice is mainly a gut-feeling.	5,00	4,00	3,00	2,00	2,00	5,00	5,00	5,00	4,00	4,00	5,00	3,00	4,00	5,00	3,00	3,00	3,00		3,82
I have based my scoring using data.	1,00	4,00	3,00	5,00	5,00	1,00	1,00	2,00	2,00	2,00	5,00	4,00	2,00	3,00	4,00	4,00	4,00		3,06
I trust that there is related information for this driver.	1,00	3,00	4,00	3,00	3,00	1,00	1,00	1,00	2,00	2,00	1,00	5,00	2,00	2,00	3,00	3,00	2,00		2,29
I can have access to this information.	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	5,00	1,00	1,00	1,00	1,00	1,00		1,24
My Trip	21,00	13,00	13,00	5,00	5,00	3,00	8,00	13,00	5,00	1,00	1,00	5,00	1,00	13,00	3,00	13,00	8,00	131,00	7,71
I have based my scoring on my knowledge (experience).	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	1,00	1,00	2,00	2,00	2,00		1,88
My choice is mainly a gut-feeling.	5,00	5,00	2,00	4,00	4,00	3,00	4,00	3,00	3,00	4,00	5,00	3,00	5,00	4,00	3,00	3,00	3,00		3,71
I have based my scoring using data.	3,00	3,00	4,00	3,00	3,00	3,00	3,00	4,00	4,00	4,00	5,00	4,00	3,00	4,00	4,00	4,00	3,00		3,59
I trust that there is related information for this driver.	2,00	2,00	3,00	4,00	4,00	2,00	3,00	3,00	4,00	4,00	1,00	5,00	3,00	3,00	3,00	4,00	2,00		3,06
I can have access to this information.	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	5,00	1,00	5,00	1,00	1,00	1,00	1,00	1,00		1,47

Figure 15: Implementation of framework in the Product Backlog of Product Owner No.6

EPIC	Number of impacted customers	Effect on user satisfaction	Loyalty - Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	External partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies on/from others	Enables business opportunities	IT Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
UserValue																			
Business value																			
Time criticality																			
Risk reduction																			
Checkin Successful (screen and option to buy)	8,00	13,00	8,00	2,00	2,00	3,00	13,00	0,00	0,00	1,00	0,00	0,00	0,00	8,00	3,00	8,00	8,00	77,00	4,53
I have based my scoring on my knowledge (experience).	1,00	1,00	1,00	3,00	3,00	3,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,41
My choice is a gut-feeling.	4,00	3,00	3,00	3,00	3,00	3,00	4,00	1,00	5,00	2,00	5,00	5,00	5,00	4,00	4,00	4,00	4,00		3,65
I have based my scoring using data.	2,00	2,00	2,00	5,00	5,00	4,00	3,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		2,18
I trust that there is related information for this driver.	1,00	1,00	1,00	4,00	4,00	3,00	2,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,59
I can have access to this information.	1,00	1,00	1,00	5,00	5,00	5,00	2,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		2,00
Go Show (checkin for earlier flight!)	8,00	13,00	13,00	2,00	2,00	5,00	1,00	0,00	0,00	1,00	0,00	0,00	0,00	2,00	1,00	1,00	2,00	51,00	3,00
I have based my scoring on my knowledge (experience).	1,00	1,00	1,00	3,00	3,00	3,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,41
My choice is mainly a gut-feeling.	4,00	3,00	4,00	3,00	3,00	2,00	2,00	1,00	2,00	2,00	5,00	5,00	5,00	4,00	4,00	4,00	4,00		3,35
I have based my scoring using data.	2,00	2,00	2,00	5,00	5,00	5,00	1,00	5,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		2,35
I trust that there is related information for this driver.	1,00	1,00	1,00	4,00	4,00	3,00	2,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,59
I can have access to this information.	1,00	1,00	1,00	5,00	5,00	5,00	2,00	5,00	5,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		2,24
New payment flow (simplification of processes)	13,00	13,00	13,00	5,00	5,00	8,00	13,00	8,00	0,00	1,00	0,00	3,00	2,00	13,00	5,00	5,00	8,00	115,00	6,76
I have based my scoring on my knowledge (experience).	1,00	1,00	1,00	2,00	2,00	2,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,24
My choice is mainly a gut-feeling.	4,00	4,00	4,00	3,00	3,00	4,00	5,00	2,00	5,00	5,00	5,00	5,00	5,00	3,00	3,00	3,00	3,00		3,88
I have based my scoring using data.	1,00	1,00	1,00	4,00	4,00	2,00	1,00	3,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,53
I trust that there is related information for this driver.	1,00	1,00	1,00	4,00	4,00	1,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,41
I can have access to this information.	1,00	1,00	1,00	4,00	4,00	1,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,41

Figure 16: Implementation of framework in the Product Backlog of Product Owner No.7

EPIC	Number of impacted customers	Effect on user/customer satisfaction	Loyalty- Repurchase	PR value	Brand value	Cost savings	Additional revenue	Contribution to Corporate Strategy	partner value	Business value decrease over time	Fixed deadline	Move to other competitor	Dependencies on/from others	Enables business opportunities	Security or privacy risk	Potential penalties & negative effects	Value of received info	Total	Average
User Value																			
Business value																			
Time criticality																			
Risk reduction																			
Redesing Paid Lounge	13	8	8	5	8	0	13	8	0	8	1	8	8	13	1	5	5	112,00	6,59
I have based my scoring on my knowledge (experience).	2	3	2	1	2	1	1	1	1	1	1	2	1	2	1	1	2		1,47
My choice is a gut-feeling.	3	2	2	1	2	5	5	2	4	5	5	2	5	3	5	5	2		3,41
I have based my scoring using data.	2	3	3	5	5	1	1	3	2	1	1	4	2	3	1	1	3		2,41
I trust that there is related information for this driver.	2	4	4	4	4	1	1	2	2	1	1	4	1	3	1	1	3		2,29
I can have access to this information.	1	4	4	4	4	1	1	3	1	1	1	4	1	3	1	1	3		2,24
Face recognition - Entrance (Future Bio-metrics usage)	8	8	1	13	8	8	1	5	5	8	1	1	5	5	0	0	13	90,00	5,29
I have based my scoring on my knowledge (experience).	2	2	5	1	2	2	1	1	1	2	1	1	1	2	1	1	2		1,65
My choice is mainly a gut-feeling.	2	3	1	4	2	2	5	4	2	2	5	5	2	1	5	5	3		3,12
I have based my scoring using data.	3	2	5	2	4	4	1	2	4	4	1	1	3	4	1	1	3		2,65
I trust that there is related information for this driver.	3	3	2	2	2	4	1	3	4	5	1	1	2	3	1	1	2		2,35
I can have access to this information.	3	3	0	3	4	4	1	3	5	0	1	1	2	4	5	5	1		3,00
Lounge Portal (WiFi Portal)	8	8	8	5	5	13	13	5	8	0	0	5	5	13	1	1	5	103,00	6,06
I have based my scoring on my knowledge (experience).	2	3	3	2	2	2	2	2	2	1	1	2	2	1	1	1	2		1,82
My choice is mainly a gut-feeling.	4	2	2	4	2	4	5	2	4	1	1	4	5	5	5	5	5		3,53
I have based my scoring using data.	2	3	3	3	3	2	1	4	2	5	5	2	4	1	1	1	2		2,59
I trust that there is related information for this driver.	2	4	4	3	4	3	2	4	2	4	4	4	2	1	2	2	3		2,94
I can have access to this information.	1	0	0	4	0	3	2	1	3	0	0	5	2	1	2	2	3		2,42

Figure 17: Implementation of framework in the Product Backlog of Product Owner No.8

5. Validation

At this chapter, we will comprehend the outcomes of the series of interviews, the results of which may be assessed in the previous chapter. It is noteworthy to mention that through the character of the semi-structured interviews that were conducted, interesting findings came up regarding the way that decision makers work and prioritize in their professional environment.

5.1 Analysis of the results – Initial Interviews

As far as the first set of interviews is concerned, we notice that half of the participants have been involved in only one project in their professional environment as Product Owners, with maximum of 5 projects in 2 different cases. The most experienced PO that was interviewed counts 3 ½ years of experience, while the average experience of all the participants as a PO is almost 1 ½ year. Out of all of the interviewees, 60% addresses their products to the end users, which would be the travelers of Airfrance-KLM. The rest 40% have internal consumers for their deliverables, which are mostly different departments than their own.

It is noteworthy to mention that almost every participant emphasized the impact of time criticality. Urgency of delivery, time of implementation/time to market and deadlines were explicitly mentioned during the interviews. Our main finding considering the concept of cost of delay was that, decision makers will often choose to deliver items that will be implemented faster, as this correlates directly to the concept of revenue, hence adding to the business value pool faster.

In that sense, business value is one the most important metrics identified during the interviews. The latter may be composed of different attributes that can enhance it beyond the economic aspect; either Brand value, or Partner and PR value, all of these may be categorized under the term of Business value.

For the purposes of this research, there has been a distinction in the terms of value when it comes to the end user or customer; nearly half of the POs that participated served other departments within the organization, without delivering their products to the customers directly. Consequently, there had to be several metrics that addressed only to the User/Customer value benefits, thus, in the implementations of these sets of metrics, the categories of Business value and User/Customer value were different.

5.2 Analysis of the results – DDDM Interviews

The Product Owners were asked to score the metrics that were introduced to them in a scale from 1-10, representing the levels of importance regarding the delivery of a software requirement. The scales were representing 1 as being of utmost importance and 10 being the least relevant to their prioritization process. Through the analysis of the responses from the 10 participants, we acquired these results (in descending order):

Metrics (Category)	**Scoring out of 10** **(Most important to least important)**
Revenue (Business Value)	1,7/10
Cost Savings (Business Value)	2,2/10
Number of impacted customers (User Value)	2,3/10
Loyalty - Repurchase* (User Value)	2,4/10
Customer Intimacy* (Non-categorized)	2,4/10
User satisfaction effect (User Value)	2,6/10
Fixed deadline (Time criticality)	2,7/10
Dependencies to/from others (Time criticality)	2,7/10
Security Risks (Risk reduction)	3,1/10
Corporate Strategy (Business Value)	3,4/10
Potential penalties/Negative effects (Risk reduction)	3,6/10
Value of received information (Risk reduction)	3,8/10
Business value reduction over time** (Time criticality)	4,1/10
New Business Opportunities (Business Value)	4,2/10
Brand Value (Business Value)	4,4/10
PR Value (Business Value)	4,7/10

Competitors Initiation (Time criticality)	5/10
Partner Value (Business Value)	5,6/10
B2E Satisfaction **(**Non-categorized**)**	5,7/10

Table 8: Perception of importance of metrics by the Product Owners

** Answered by 5 Product Owners instead of 10.*
***Answered by 9 Product Owners instead of 10.*

Going through the numbers, we realize that Revenue is considered the most important metric, scoring 1,7 out of 10 and being categorized between the scale 1 and 2. This result was more or less anticipated, as the decision makers correlate the results of their work to the immediate return of income for the organization, in spite of the fact that not all of the requirements delivered aimed directly the customer. Furthermore, marked as highly important metrics, mentioned in a priority sequence, were cost savings and the number of impacted customers, scoring 2,2 and 2,3 out of 10 respectively. POs stressed out the importance of saving money as a result of changing their Product Backlogs, delivering these items as soon as possible. As for the customers impacted, all of the POs regardless of delivering their products directly to the end users or not, highlighted the need to deliver products that would immediately improve the customers perceptions about the product of the company.

Moving on, the metrics that follow in this prioritized list are the repurchase tension of the customer, as well as customer intimacy. It is interesting to notice that only half of the participants provided answers about these metrics; this is more due to the fact that their deliverables aimed stakeholders internal to the organization, rather than the end user, thus their response would not be objective. Nevertheless, from the responses that we acquired, these two metrics scored as very important, between the scale of 2 and 3, and more specifically 2,4 out of 10. We must note at this point that, due to the correlation of the nature of the metric “Customer intimacy” with the metric “User satisfaction”, and due to the reduced amount of responses that we obtained for the former, we applied only one metric out of those two in the final implementation of the framework; the metric that was evaluated was “User satisfaction effect”. The latter, which applies to both internal users as well as to the end user, scored 2,6 out of 10, and fixed deadline as well as the fact of dependencies from other teams scored 2,7 out of 10 respectively.

Furthermore, marked as important, in between the scale of 3 and 4, we find security risks, scoring 3,1 out of 10. It goes without saying that decision makers always take into account probable faults that may lead to vulnerable situations regarding the product, that need to be eliminated as soon as possible. Next to this scale comes the alignment with the corporate strategy; items delivered that comply with the company’s mission and vision are taken into consideration more than others, proved by the scoring of this metric with 3,4 out of 10. The last two metrics in this scale are the potential penalties or negative effects that might occur by not delivering the specific item or by violating a law or an agreement, as well as the value of received information, which refers to the addition of the knowledge pool of decision makers for future consideration and retrospection, scoring 3,6 out of 10 and 3,8 out of 10 respectively.

In between the next scale, marked as fairly important, we came across four metrics. Business value reduction over time refers to the loss of value over time while the company fails to deliver the item, scoring 4,1 out of 10. It is noteworthy that this metric also averages to this amount by the responses of 9 POs instead of 10. For the rest metrics, we have new business opportunities, for which the company capitalizes on changing trends or increasing demands for a product, scoring 4,2 out of 10. This is followed by the Brand value of the organization, scoring 4,4 out of 10, and PR value which refers to the practice of creating goodwill of the organization's publics, such as employees and investors through publicity, scoring 4,7 out of 10.

Lastly, marked as of neutral importance, we come across to the metrics scoring in between the scale of 5 and 6; competitor's initiation does not seem to trouble the decision makers that much, which refers to the probability of the scenario that the competitors move towards a specific solution before us, scoring a net 5 out 10. The rest metrics in this bracket are Partner value, which refers to the value that shareholders obtain from their investments in the company, scoring 5,6 out of 10, and B2E satisfaction, which addresses the levels of satisfaction achieved regarding the employees of the organization, scoring 5,7 out of 10. At this point we need to note, that due to the fact that the majority of the decision makers did not evaluate this metric as important, and due to the fact that it scores last to the list of our metrics that we asked about, we decided to not include it in the final framework.

In the next section of questions, POs we asked to evaluate several statements regarding their perception about their prioritization process.

Statements	**Scoring out of 5 (From highest to lowest)**
Belief that data driven practices may improve business outcomes	1,3/5
Belief that transparency of prioritization to stakeholders can be enhanced	1,7/5
Willingness to apply new methods to standardize prioritization procedure	1,9/5
Personal technical fluency	2,3/5
Understanding of DDDM principles and practices	2,4/5
Confidence for current prioritization procedure	2,7/5

Satisfaction of prioritization transparency to stakeholders	3,3/5

Table 9: Perceptions of the Product Owners about the prioritization process

Our main takeaways from these inquiries come down to the decision makers' strong belief that data driven practices may only enhance the quality of the business processes and outcomes. They also strongly agree that the transparency of their practices to their stakeholders may be enhanced. This is a very important finding, as POs have to communicate many times to their stakeholders the reasoning behind their decisions, which may also lead to conflicts and misinterpretations, if not performed properly. For that matter, they are willing to try out new methods that will assist them automate their prioritization procedure and to avoid explaining to every person of interest separately why they deliver a specific item over another. The POs rate their personal technology fluency, as well as their understanding of DDDM principles fairly high, but they all conclude to the fact that they may enhance their knowledge pool in order to be more data driven oriented. These soft skills may be boosted with better communication amongst the different departments within the organization. This would eventually lead to improve their self-esteem regarding the prioritization procedures that they follow.

Lastly, another interesting finding is that the current levels of satisfaction regarding the communication of their decisions to their stakeholders are not that high, which means that they would appreciate a way to standardize their processes in order to prove with more evidence why they pick several deliverables over others.

In the last section of this interview, we tried to educate the POs over the recognized levels of DDDM according to Jia et. al, 2005, and to ask several questions regarding each of the three tiers within it, addressing to the collection of data, the analysis of data, and the synthesis of data.

As far as the collection of data is concerned, we collected the following responses:

Statements	**Scoring out of 5 (Strongest belief to weakest belief)**
Collecting data is time consuming	2,4/5
Data collection methodologies are being conducted properly*	2,5/5
As a department, we are dependent for important customer's data from other departments	2,5/5

As a department, we have open discussions about data.	2,7/5
As a department, we collect data for all of our customer's segments	2,87/5
As a department, we have sufficient resources to collect data	3/5
As a department, we have an efficient data collection system in place	3,2/5
As a department, we have access to preferred datasets easily	3,5/5
As a department, we archive all the data that we receive properly*	3,67/5

Table 10: Perceptions of the Product Owners about data collection

** 4 POs answered Do not know.*

The most prevalent statement is that decision makers consider the process of collecting data as time consuming. It only makes sense to score approximately the same with the statement that their departments are dependent on important customer's data from other departments. In that sense, when the stakeholders increase and the sources that one may obtain valuable information multiply, then the procedure of collecting all of the important data is being prolonged. In the same scoring levels we find the statement that data collection methodologies are being conducted properly within the departments of the participants, though it is interesting to note that only 6 out of 10 participants answered this statement.

In the same scale, in which the participants all agree, we come across the facts that POs have open discussions about data and that they collect data for various customer segmentations.

The participants evaluated as neutral the statement that they have sufficient resources to collect data. This is due to the fact that representatives from several departments addressed the issue of not being able to access data of several customer segments that they need, and that they have to request the implementation of new entities in order to acquire these datasets, which is a timely procedure. This explains the fact that it scored almost the same with the statement that there is an efficient data collection system in place.

Lastly, the majority of the responses concludes to the fact that they do not have access to their preferred datasets easily, a statement that is rationalized by our previous findings. It is interesting to note here that the participants feel that there is not a proper methodology for archiving the data that they receive in a desirable manner, although only 6 out of the 10 participants scored this statement.

As far as the analysis of data is concerned, we collected the following responses:

Statements	Scoring out of 5 (Strongest belief to weakest belief)
Analysis improves services we provide to customers	1,2/5
Analysis improves the skills and job performance of employees	1,7/5
As a department, we have staff with expertise in data analysis	2,8/5
Translating data takes away from the time spent collecting data *	3,2/5
As a department, we translate discussions of data into actions	3,5/5
I have received training to effectively interpret and act upon results	3,8/5

Table 11: Perceptions of the Product Owners about data analysis

** 5 POs answered Do not know.*

In this series of questions, the responses are a little bit clearer. With almost an absolute consensus, POs strongly agree that the analysis of the data acquired improves the services that are provided to the customers. In the same scale, the participants feel that the analysis may improve the skillsets of the employees, which leads to a better job performance and a more efficient deliverable final product.

A statement that scores close to the neutral scale perception is that the department of each participant has staff with expertise in data analysis. POs feel that this is mainly a responsibility of the Digital Analytics department, thus they are not aware of people responsible for this kind of practices within their domain.

Another statement that scores almost the medium average is that translating data is taking away time from time that could be spent in collecting data. This statement cannot provide us with any reasonable insights, as only half of the participants provided a scoring, indicating that the participants are not aware if this statement stands or not.

Lastly, some interesting findings are that the departments of the participants take limited actions to translate discussions about data into real implementations. This may be justified due to the fact that the insights need to be communicated to the according stakeholders, and that it usually occurs only when a necessity about a customer's data segmentation is acknowledged. This finding may also be justified by the fact that the participants do not feel adequately literate to interpret and act upon results.

As far as the synthesis of data is concerned, we collected the following responses:

Statements	**Scoring out of 5 (Strongest belief to weakest belief)**
Our main focus is to implement our features and deliver on time	1,5/5
Our main focus is to increase customer satisfaction and quality of service	1,8/5
Our main focus is to increase the business value constantly	2/5
We acquire sets of information from multiple sources	2/5
Using data builds an understanding of how our domain operates	2,2/5
Our main focus is to reduce potential risks for our domain and organization	2,9/5
As a department, we use data to improve the quality of decision-making	3,1/5
As a department, we use data to reflect on our decisions	3,6/5

Table 12: Perceptions of the Product Owners about data synthesis

The participants strongly agree that their main focus is to implement what is due and deliver on time. Through that, they aim to steadily increase Quality of Service (QOS) as well as customer satisfaction.

In a smaller scale of importance, POs agree that they focus their aim to have the business value increased, and mostly interpreted in terms of revenue. Accordingly, they also agree that they acquire sets of information for different segments of the customer from multiple sources and that using all of this data, translated into information, helps them understand better how their domain operates.

Lastly, the participants hold a neutral stance when they were asked if they aim to reduce potential risks for the organization. This leads us to understand that it is not one of their main priorities. Furthermore, they do not claim that they use data to improve the quality of the decision making process, but they do not deny using it when it is available. It is clear though, that they use much less data as a component of their retrospection regarding the outcomes of their decisions.

5.3 Analysis of the results – Product Owner Matrix Interviews

Considering the implementation of the sets of metrics in the personal Product Backlog of each Product Owner, we asked them to answer several statements while they were applying the Fibonacci enumerations regarding each metric. These statements will assist us understand how much they base their decision on datasets that are available, if their answers come from educated guesses or from their experience, and if they may have access to information related to each specific metric, if any. In this part, we will address to these sets of metrics as the PO matrix, and we shall realize if this implementation agrees with their current prioritization.

The names of the epics will be presented as they were provided by the Product Owners at that moment. The averages of the metrics will be presented with a 2 digit decimal number, in order to make the prioritization of each item clearer. In the end, we add up all of these evaluated metrics into a total sum, and then using the average of these numbers, we will understand which item needs to be delivered first, taking under consideration the one that scores higher. In total, 17 metrics were taken under consideration.

As far as the statements are concerned, we will present the averages of the responses for each epic, and later on we will provide the average number of their responses in order to understand their perception about the process. While calculating the averages, when the respondent replied with 0, standing for 'Do not know', his response was not calculated in the medium average of the evaluated statements.

It needs to be stressed out that the Fibonacci sequence is based on the enumeration: (**0,1,2,3,5,8,13,21**). A Product Owner scores a metric according to its importance following the Fibonacci sequence, marking with '0' a metric as one to be non-important and with '21' a metric that is very important for the current Epic that is under investigation. For instance, the PO scores '13' regarding the metric 'Number of impacted customers'; this would imply that the implementation of this software requirement would immediately affect a big number of customers. On the other hand, if a metric such as 'Brand Value' scores '1', it would mean that the PO trusts that implementation of this software requirement would have minimal impact on the contribution to the company's Brand value.

Further on, the charts that follow each figure are representing two different concepts. The first chart under each figure depicts the averages that each Epic scores, showing us that the ones scoring the biggest averages need to implemented over the others. The second chart is listing 5 statements that have to do with the scoring of each Epic. An average of the responses for each statement is formed in the last column, which is enumerated in a scale from 1 to 5, with '1' being 'Strongly Agree' and '5' being 'Strongly Disagree'.

For the results of the Product Owner No.1 :

The Product Backlog contained 4 epics that needed prioritization.

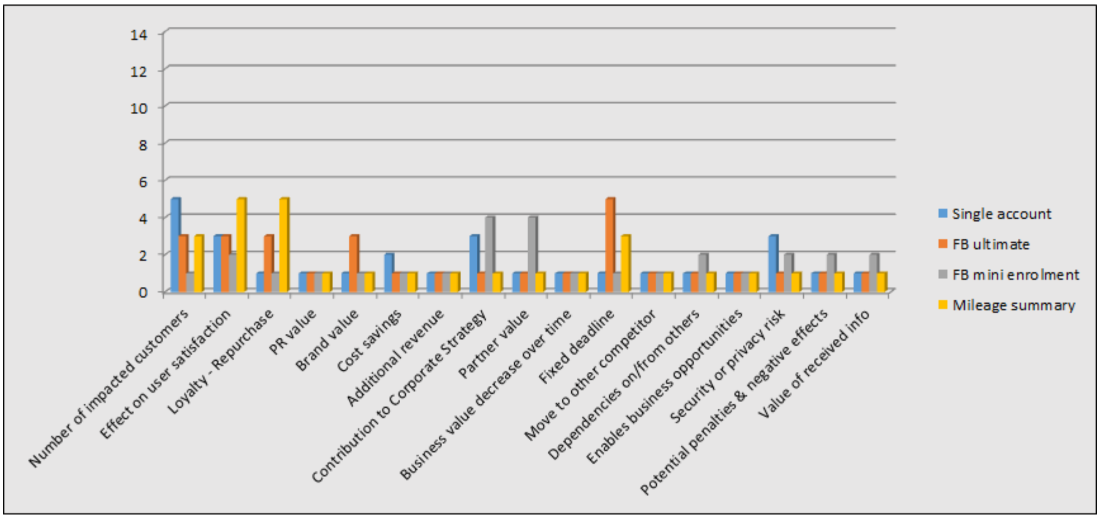

Figure 18: Scoring of each epic on every given metric for Product Owner No.1

Epic	Average of metrics
Single account	1,65
FB ultimate	1,71
FB mini enrolment	1,65
Mileage summary	1,71

Statement	Epic No1	Epic No2	Epic No3	Epic No4	Average
I have based my scoring on my knowledge.	4	4	4	4	4
My choice is a gut-feeling.	4	3	2	4	3,25
I have based my scoring using data.	2	4	4	2	3
I trust that there is related information for this driver.	3	4	4	3	3,5
I can have access to this information.	2	3	2	2	2,25

In this case, the PO provided unified numbers in his responses. Considering the prioritization of his epics, the numbers are really close, making it hard to evaluate efficiently which items need to be prioritized over others. Nevertheless, the PO responded that indeed the items that scored 1,71 were already higher in his list, and the rest two were following. For the prioritization amongst the ties, he took under consideration the feasibility of implementation at this particular point and which item may be delivered faster.

In general, the PO informs us that this procedure was hardly based on his knowledge about the metrics. Averagely, he uses some sort of data in order to inform his decision, and in the same manner, he trusts his instinct in order to score the metric. His perception about the availability of related information regarding the metrics is that there is a possibility that there is not any available, but he trusts that he can access this information if it is available.

For the results of the Product Owner No.2 :

For the prioritization, the Product Owner picked the following 2 epics.

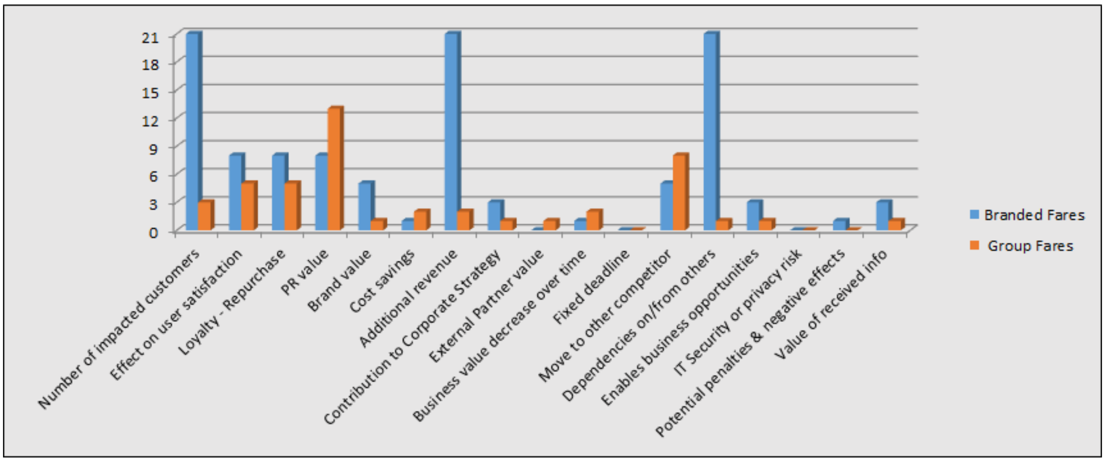

Figure 19: Scoring of each epic on every given metric for Product Owner No.2

Epic	Average of metrics
Branded Fares	6,41
Group Fares	2,71

Statement	Epic No1	Epic No2	Average
I have based my scoring on my knowledge.	2,35	1,94	2,14
My choice is a gut-feeling.	3,29	3,88	3,58

I have based my scoring using data.	3,35	3,82	3,58
I trust that there is related information for this driver.	2,5	2,08	2,29
I can have access to this information.	2,13	1,06	1,59

From the beginning, it was clear for the PO that Branded fares, which is a depiction of different prices to the customer, was more important to be implemented than Group fares, which refers to offering a fare to a group. The majority of the metrics got a high mark in the Fibonacci sequence regarding the former, as this was not the case for the latter. This explains the difference between their averages, and through this implementation we verified that the current prioritization was valid.

It is interesting to note that the participant was not aware of any related information of several metrics to his epics. These metrics were “Loyalty-Repurchase”, “Brand Value” and “Competitor’s Initiation”. In total, for Branded fares, he declared that he was not aware if there is related information for 5 metrics, for two of which he did not know if he can even access this information. For Group fares, there were 4 metrics noted in total, that he was not aware of any information related to them.

In this case, the participant felt pretty confident about his knowledge, basing the scoring of the metrics mostly on his experience rather than intuition. The scoring did not come though by going through any sort of data, he trusts that there is related information for every metric that is mentioned and that he can have access easily to each one.

For the results of the Product Owner No.3:

For the prioritization, this Product Owner used 3 epics from his Product Backlog:

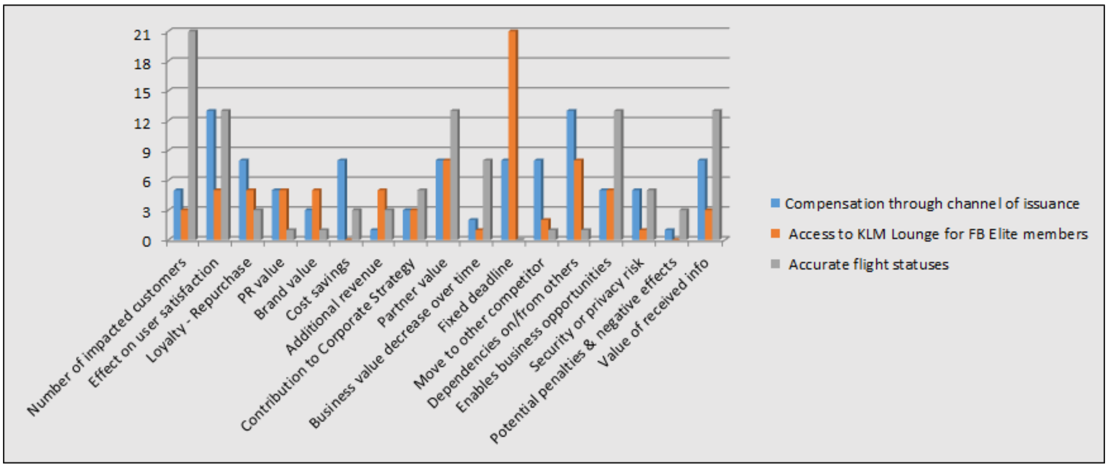

Figure 20: Scoring of each epic on every given metric for Product Owner No.3

Epic	Average of metrics
Compensation through channel of issuance	6,12
Access to KLM Lounge for FB Elite members	4,71
Accurate flight statuses	6,29

Statement	Epic No1	Epic No2	Epic No3	Average
I have based my scoring on my knowledge.	3,29	4,18	3	3,49
My choice is a gut-feeling.	2,35	1,82	3	2,39
I have based my scoring using data.	5	4,76	3,88	4,54
I trust that there is related information for this driver.	2,21	2,33	1,38	1,97
I can have access to this information.	2,64	2,22	1,69	2,18

After the prioritization ended, the participant mentioned that, through this process, it was now clear which item needs to be pursued in order to be implemented first. Accurate flight statuses came up to be the most important item for delivery, followed by compensation through channel of issuance, as well as access to KLM Lounge for FB Elite members. Since the epics were not prioritized yet in the PO's Backlog at that point, we concluded that this was a helpful session in order to provide some insights to him and move on accordingly.

An interesting observation is that the participant mentioned that he could not have access to information out of 8 metrics mentioned, considering the evaluation of his second epic. This finding can tell us that not every epic may be informed properly from these metrics, as there might not be related information.

In general, the participant mentions that he did not base his answers on his knowledge, rather his ratings for the metrics leaned to be more of a gut feeling. It is clear that he did not use any kind of data to determine his scoring, but he trusts that there is relevant information for each metric and that it will be easy for him to access it.

For the results of the Product Owner No.4:

During this implementation, the PO used 3 epics from his Product Backlog.

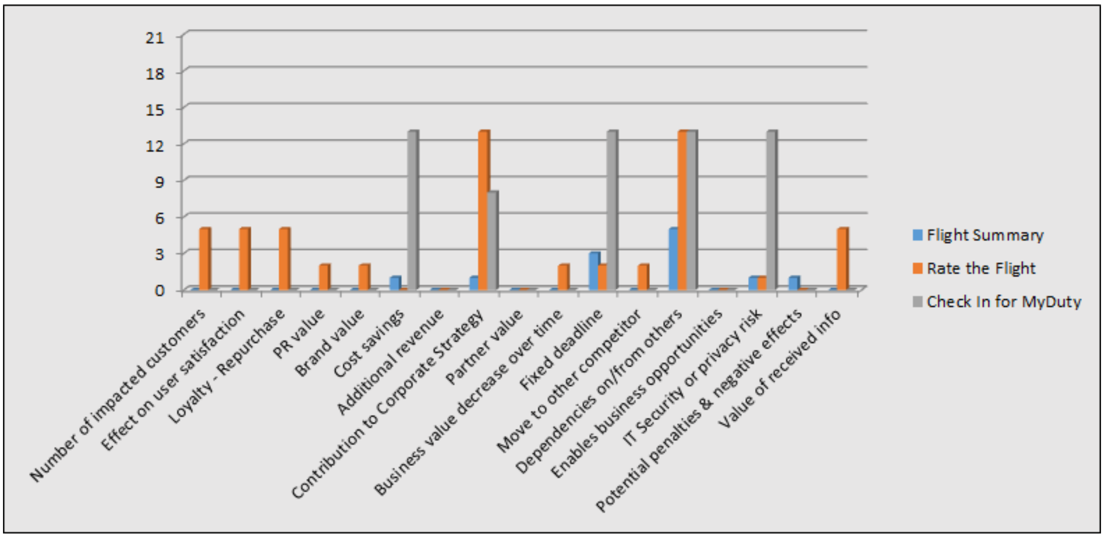

Figure 21: Scoring of each epic on every given metric for Product Owner No.4

Epic	Average of metrics
Flight Summary	0,71
Rate the Flight	3,35
Check In for MyDuty	3,53

Statement	Epic No1	Epic No2	Epic No3	Average
I have based my scoring on my knowledge.	2,76	2,94	1,647	2,45
My choice is a gut-feeling.	1,29	1,53	2,059	1,63
I have based my scoring using data.	5	5,00	4,765	4,92
I trust that there is related information for this driver.	3,73	2,57	1,67	2,65
I can have access to this information.	3,73	1,83	1,67	2,41

Two out of three items were significantly more important to be delivered in comparison to the other one (Flight Summary). ‘Rate the Flight’ and ‘Check In for MyDuty’ are pretty close regarding their scoring, with the latter one ranking higher, thus needed to be implemented first. Indeed, the participant agreed that eventually the 'Check In' epic needs to be delivered first, and he agreed that this would be the proper way to prioritize the items of his Backlog.

An interesting observation during this session is that the participant claimed to score all of the metrics based on intuition and definitely not on data. As we can see, one of the most prevalent statements is that his choices were mainly gut feelings, and almost nothing was based on data. He trusts that there is related information about these drivers and that he could access them, but his scoring was basically based on his knowledge more or less and most certainly on educated guesses.

This of course happens only when he trusts that there is related information available. For that matter, it is important to note that for the two most important epics, the participant was not aware if there is related information and if he could access it regarding 10 metrics for the 'Rate the flight' epic and 11 metrics for the 'Check In' epic. It is clear at this point, that the particular participant wants this process to be more standardized and based on facts that may assist him, rather than basing the whole prioritization process onto a gut feeling.

For the results of the Product Owner No.5:

The specific participant brought 3 of his epics from his Product Backlog to the table.

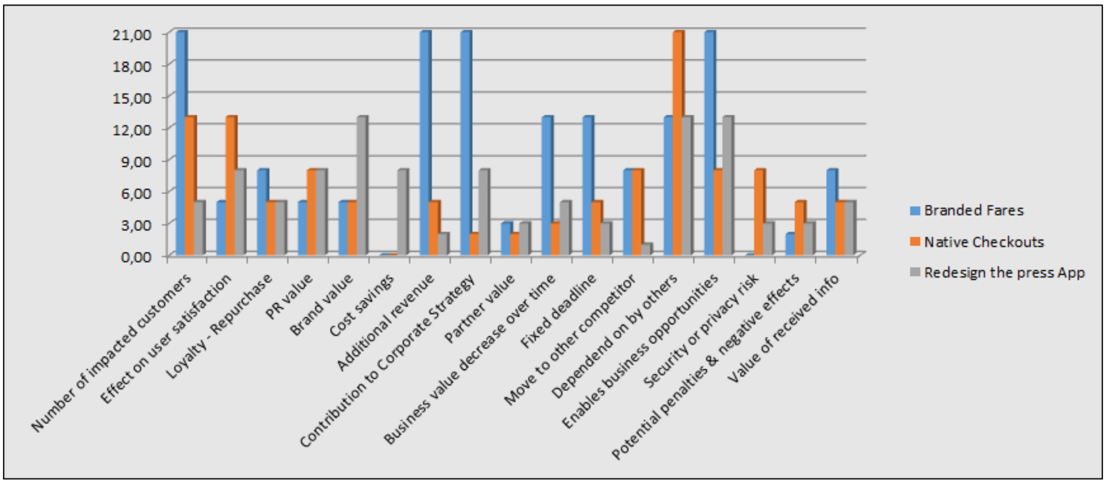

Figure 22: Scoring of each epic on every given metric for Product Owner No.5

Epic	Average of metrics
Branded Fares	9,82
Native Checkouts	6,82
Redesign the press App	6,24

Statement	Epic No1	Epic No2	Epic No3	Average

I have based my scoring on my knowledge.	3,29	3,76	3,47	3,5
My choice is a gut-feeling.	1,65	1,41	2,53	1,86
I have based my scoring using data.	5,00	4,94	4,88	4,94
I trust that there is related information for this driver.	3,88	4,29	3,35	3,84
I can have access to this information.	4,18	4,12	3,12	3,8

After this implementation, the participant confirmed that it was clear to him which epics need to be implemented first. Indeed, 'Branded Fares' was already ranked allegedly as first in this Product Backlog, but now the reasons hindering behind his choice were clear to him as well. The rest two epics scored pretty much close averages, yet again it was clear that Check-Outs needed to be implemented first; at that given point, the PO had not made up his mind regarding which item needed to be delivered first, so he confirmed that this procedure was insightful and that he would definitely take it under consideration to rank his Backlog in a similar way to the outcome of this interview.

It was evident throughout the whole process that the participant did not base his scoring on data at all. Genuinely, his evaluations about the metrics were clearly a gut feeling and less based on his experience or knowledge on each subject.

An interesting finding in this interview was the fact that the participant trusted that there are not available information about these metrics at the moment, and even though he had this information, he would not be able to access it, making him incapable of evaluating the metrics reasonably.

For the results of the Product Owner No.6:

The participant decided to evaluate three epics from his vault, as he had quite few items that needed to be prioritized, but this whole process would take more time than scheduled in order to evaluate all of his Product Backlog items.

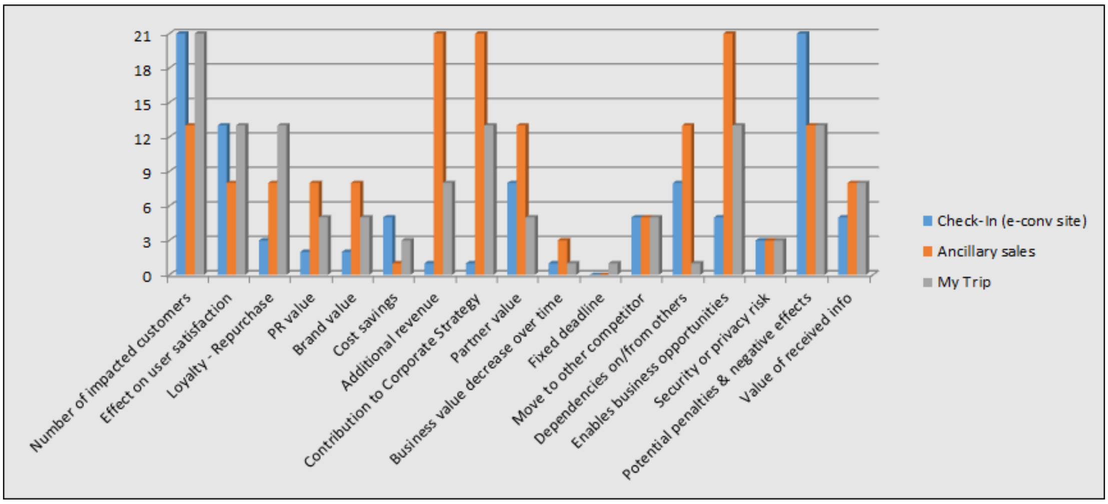

Figure 23: Scoring of each epic on every given metric for Product Owner No.6

Epic	Average of metrics
Check-In (e-conv site)	6,12
Ancillary sales	9,82
My Trip	7,71

Statement	Epic No1	Epic No2	Epic No3	Average
I have based my scoring on my knowledge.	2	2,18	1,88	2,02
My choice is a gut-feeling.	3,41	3,82	3,70	3,64
I have based my scoring using data.	3,12	3,06	3,59	3,25
I trust that there is related information for this driver.	2,70	2,29	3,06	2,68
I can have access to this information.	2	1,24	1,47	1,57

When the implementation of the PO Matrix framework was concluded, it was evident that Ancillary Sales had to be prioritized over the 'My Trip' epic, which was ranked higher than

Check-In. The participant had not ranked the items in the way the results turned out, and it was of real value to him to have this transparent procedure in hand, in order to re-prioritize if needed. An interesting fact is that during this process, the participant evaluated the metrics that found interesting with rather high marks, which resulted into the big final average numbers.

During the evaluation of the scoring procedure, the PO concluded that he neither based his scoring on gut-feeling, nor using thorough analysis of data. His scoring was the result of his experience on the domain, and on his knowledge about the metrics, as he perceived them. He was more confident that there is related information for these metrics and that he could acquire it in case he wanted it, noticing that he would be able to access it easily if needed.

For the results of the Product Owner No.7:

For this implementation, the participant also wished to use 3 items from his Product Backlog, due to time constraints.

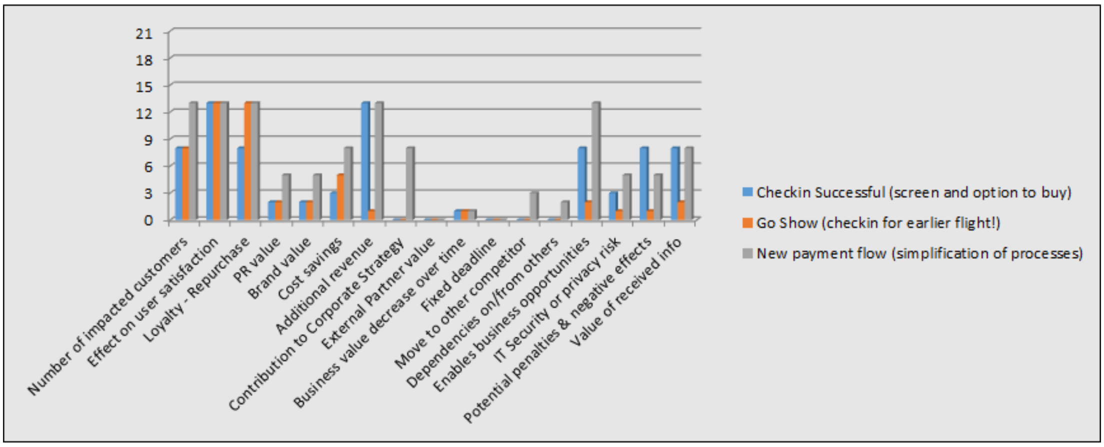

Figure 24: Scoring of each epic on every given metric for Product Owner No.7

Epic	Average of metrics
Check-in Successful (screen and option to buy)	4,53
Go Show (checkin for earlier flight)	3,00
New payment flow (simplification of processes)	6,76

Statement	Epic No1	Epic No2	Epic No3	Average
I have based my scoring on my knowledge.	1,41	1,41	1,23	1,35
My choice is a gut-feeling.	3,65	3,35	3,88	3,62

I have based my scoring using data.	2,18	2,35	1,53	2,02
I trust that there is related information for this driver.	1,59	1,59	1,41	1,53
I can have access to this information.	2	2,24	1,41	1,88

In this case, the participant had a clearer picture of his implementation and seemed confident about the way that he scored the metrics. The outcome was of not a surprise to him, as the results matched to the prioritized list that he already had. It came out eventually that it was of utmost importance to have the New payment flow implemented first, which he had already under implementation, followed by the Check In and the Go Flow, which were meant to be implemented in a later stage.

The PO based his scoring mostly on his knowledge about the metrics. His experience worked to his advantage regarding the evaluation of each metric, and he claimed to have used efficient amounts of data in order to deliver his evaluations. His choices were not a product of intuition; he strongly trusted that there is available information about each metric and that it would be fairly easy for him to access it if he wanted to.

For the results of the Product Owner No.8:

The last participant also thought wise to use three items from his personal Product Backlog, in order to confirm his current prioritization regarding these epics.

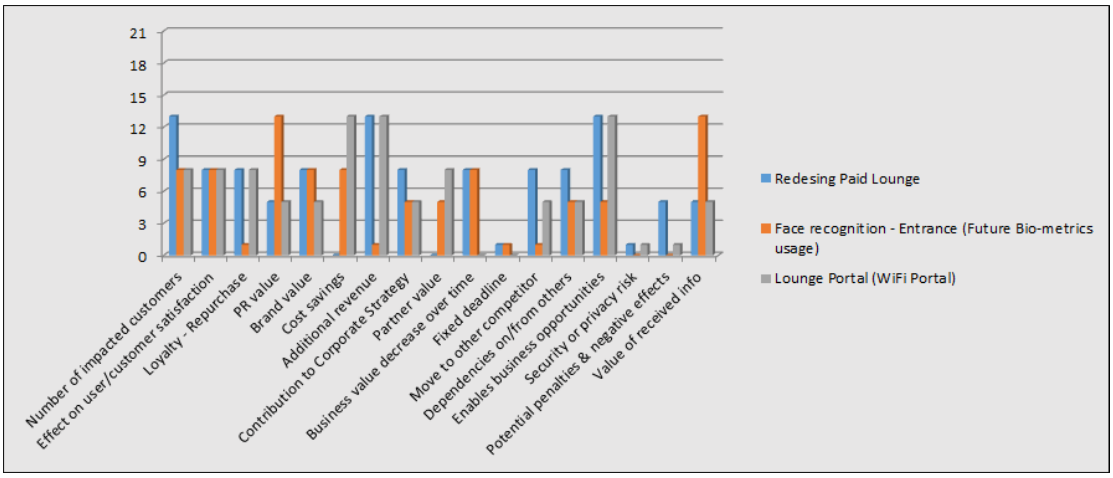

Figure 25: Scoring of each epic on every given metric for Product Owner No.8

Epic	Average of metrics
Re-desing Paid Lounge	6,59
Face recognition - Entrance (Future Bio-metrics usage)	5,29

Lounge Portal (WiFi Portal)	6,06

Statement	Epic No1	Epic No2	Epic No3	Average
I have based my scoring on my knowledge.	1,47	1,65	1,82	1,64
My choice is a gut-feeling.	3,41	3,12	3,53	3,35
I have based my scoring using data.	2,41	2,65	2,59	2,55
I trust that there is related information for this driver.	2,29	2,35	2,94	2,52
I can have access to this information.	2,24	3,00	2,42	2,55

After the conclusion of the evaluation of the set of metrics, the participant witnessed that the average outcome of the three epics was fairly scored evenly. It was evident that re-designing the Paid Lounge was more important than implementing the Wi-Fi Portal, a finding that did not agree with his current prioritization. This outcome is logical, since the outcomes of these two score very close; nevertheless, he trusted that for the time being he would stick to his current prioritized list and that he would not adapt to the new results, as he had the latter epic ongoing. It was mutually agreed that the Face Recognition epic would have to be implemented in a later stage.

In this case also, the participant made use of his knowledge on his domain, basing his scoring on his experience mostly. He stated that he made use of data in order to deliver the evaluations, but not extensively, and that his outcomes were less of a result of gut feelings. Finally, he trusted that with some research and by addressing to the right stakeholders, he could find all the related information about the metrics that he needed, and that the information itself may be found but not in the most easily accessible way.

6. Conclusions

6.1 Discussion

The purpose of this research thesis was to conceptualize and validate a set of metrics that may be educated through data, in the context of a case study within KLM - Royal Dutch Airlines. The two-folded character of this research was to identify the literature gaps that were recognized through a Systematic Literature Review of a number of articles and to implement a proposed framework of metrics applied in the Product Backlogs of Product Owners. The latter are dealing with delivery of software requirements in an agile setup, in order to evaluate how much of assistance would it be for them to inform their decisions in a more data driven manner.

Initially, the conceptualization of this set of metrics occurred from interviews with decision makers within the organization; Product Owners, who are responsible prioritizing software requirements, identified the need to educate their processing based on metrics and data. From the introductory interviews, we gathered audio files for over 6 hours and transcribed material that summed up to over 17300 words. With this knowledge, the finalized set of metrics under investigation started to take its form, as we were able to identify the sectors that decision makers found interesting to educate their prioritization upon.

From the second sets of interviews, we confirmed that Product Owners seek to make their justifications more data driven. The idea, according to their responses, is to decide based on numerical evidence in several occasions; in the form of revenue reports or the number of clicks in a page, or on qualitative data, such as customer suggestions/feedback, complaints from social media or in terms of operational damage, in case something goes wrong with a passenger. These interviews showed us that a very little number of measurable metrics can be used within the organization in order to frame a holistic approach about a software requirement. This is realized through material that resulted from transcribing over 4000 words from audio files, as well as quantitative material from scoring statements, that assisted us to understand which common perceptions prevail within the organization.

Eventually, the validation of this set of metrics occurred with the proposed framework that we put into use. The participants evaluated each metric in each category, in spite of the fact that they considered several metrics to be way more important than others regarding the delivery of their items. In this sense, we may now propose some other ideas for future research, regarding the weighing of these metrics, a concept that is explained further on. In total, 7 out of 8 decision makers that participated in the last series of interviews came to the conclusion that this proposed framework could be of use to them, if they had proper access to information related to these metrics.

It is noteworthy to mention that this procedure lasted one hour in average, regarding application in three (3) items for each Product Backlog, in most of the cases. We can safely notice that his procedure is timely, and many of the participants were reluctant in the beginning about its functionality. Eventually, each participant concluded that, after the second implementation, the evaluations of the metrics would be much quicker, and since they would not have to elaborate on answering the statements that were part of the current research, they believe that it could be a product of an automated process in the future. It will be of utmost value to provide a document in a form of a sheet, that will be able to inform all of the stakeholders the reasons why several items are delivered over others. The amount of time that will be saved concerning educating and discussing with the stakeholders will be much more, in comparison to the extra time that the POs will need in order to implement the framework for each item in their Backlog.

As regards the first sub-question of our main research question, we conclude, from the narrow perspective of these interviews' results, that decision makers would mainly consider the direct benefits for the company in terms of revenue and cost savings, thus business value. Metrics within the bracket of business value and user value are appreciated more when it comes to decide upon a requirement, and this can be realized by Table 8 in section 5.1.2, in which we score the average responses regarding the importance of each one. Through the interviews, we realized that each decision maker perceives a metric as a quantifiable measurement, that its value may be used for numerical comparison. The amount of crashes, user traffic, time spent in the booking API, number of error pages are perceived as metrics from the decision makers; rough numbers that depict the functionality of an implementation. This makes sense if we realize that, for core business functionalities, Product Owners do not need to apply DDDM, because hygiene software requirements are crucial for the functionality of the applications. Our main finding is that, if they mean to orient their decisions in a broader perspective, they need to have access to metrics that are beyond rough data.

It is interesting to note though, that not much attention is given to the metrics that cannot be educated through quantitative data, such as Partner Value, PR Value and Brand Value, or even New Business Opportunities. In that sense, we are able to answer to the second sub-question of this thesis research, as regards to how can the decision makers be more data driven. It is noteworthy to mention that Product Owners appreciate the possibility to evaluate more metrics, but they are limited due to the fact that they cannot find relevant information about them. We will try to elaborate on several findings that assist us to perceive the metrics better.

First of all, we will take a look at Public Relations (PR) value. According to Macnamara, "PR and corporate communication programs need to have S.M.A.R.T. objectives – objectives that are specific, measurable, achievable, relevant and timely – and which are also aligned with the over-arching objectives of the organization. An approach which allows specific objectives to measure the direct impact of PR and corporate communication as well its longer-term contribution to overall organizational objectives is micro measuring and macro measuring in two stages. Specific objectives of PR and corporate communication must be agreed with management – management needs to 'buy in' to PR objectives, recognizing them as contributing to the overall objectives of the organization. Macro-measuring refers to measuring over-arching organizational outcomes against desired objectives. Micro-measuring refers to the determination of the results of specific communication activities such as events, product launches, media publicity, analyst briefings, etc. While measurement at the macro level is ultimately the most important, micro-measuring to establish the effects, if any, from specific communication activities is important (a) to determine their success or otherwise and whether they should be continued, discontinued or changed and (b) to progressively track cumulative contributions to overall outcomes, as communication is usually a long-term multi-faceted task."

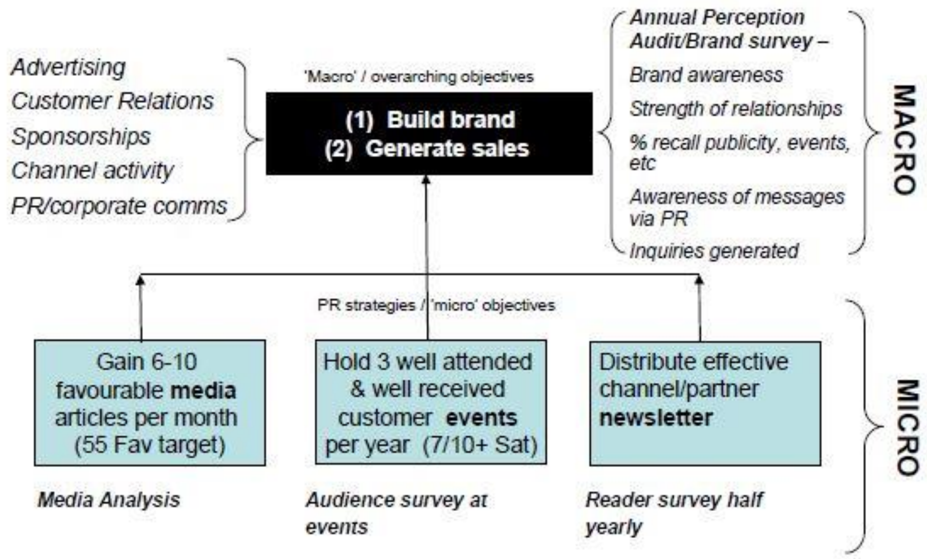

Figure 26: Micro and macro measurement of PR value, Macnamara (2015)

Based on figure 26, Macnamara states that "to ensure alignment of PR and corporate communication objectives and recognition of this alignment, it is highly recommended that we discuss objectives with management and ask the question: "If we achieve X and Y, do you agree that it will contribute to the organization's key overall objectives?" If management cannot see that what you are doing or proposing links to their overall objectives, you should not proceed. Either education of management or realignment of the objectives is required." The main takeaway from his paper is that PR practitioners need to learn to talk the language of management – numbers, percentages, charts and graphs – and express the outcomes of their work in those terms.

In his paper, Macnamara describes several best practice models for PR research; we will address to The Macro Model of PR Evaluation, renamed the Pyramid Model of PR Research (Macnamara, 1992; 1999; 2002), in which a large number of research and evaluation methodologies available to practitioners is described. Specifically, these are:

- Secondary data (ie. existing research) which can be accessed within the organization (e.g. Market research, employee surveys, customer complaints data, etc.) or externally from the Web, the media, research services such as Lexis-Nexis, academic journals etc;
- Advisory or consultative groups;
- Online 'chat rooms' and other informal feedback mechanisms;
- Unstructured and semi-structured interviews;
- Readability tests on copy (e.g. Fog Index);
- Pre-testing (e.g. PDF files of proposed publications, mock-ups of Web pages);
- Response mechanisms such as Web statistics.

As far as Partner Value is concerned, Banford and Ernst[2] (2002) inform us through their work, with more than 500 companies around the world involved, that fewer than one in four

[2] http://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/measuring-alliance-performance

alliances have adequate performance metrics in place, so companies need to develop a more structured approach to evaluate the health of their alliance. Impediments to measure the alliance performance are at first the difficulty to agree on a single measure of performance; secondly, the operational interdependence makes benefits and costs difficult to track because complicated transfer-pricing issues are created, and lastly the noncore position of alliances within the corporate portfolio prevents them from receiving the same level of management scrutiny as business units. A balanced view of performance can be achieved by including four dimensions of performance fitness; financial, strategic, operational and relationship. We will present the related metrics according to the article:

- ***Financial fitness***: Sales revenue, Cash flow, Net income, Return On Investment, Expected net present value of the alliance, partner specific metrics (transfer-pricing revenues, sales of related products)
- ***Strategic fitness***: Market share, New product launches, customer loyalty, competitive positioning
- ***Operational fitness***: Number of customers visited, staff members recruited, quality of products, manufacturing throughput
- ***Relationship fitness***: Cultural fit and trust between partners, speed and clarity of decision making, effectiveness of interventions when problems arise.

Lastly, another metric that needs to be more assessed is Brand Value[3]. A sound instrument to measure brand value is not in place. It may include trademarks, logos, packaging, market strategy, digital assets, brand colors etc.; really anything that customers can associate with brand image. Yet again, a few valuation methods could include:

- ***Cost based brand valuation***: The sum of individual costs or values of brand assets and liabilities. It is the accumulation of costs that incurred to build the brand since its inception, and it could include historical advertising, promotion expenditures, costs of campaigns, licensing and registration costs.

- ***Market based brand valuation***: It addresses to a comparison of similar brands that have been sold, using comparable market transactions, like the specific sale of a brand or stock market quotations.

- ***Income approach brand valuation***: It considers the valuation of future net earnings that directly attribute to the brand, to determine the value of the brand in its current use. The brand value using this method is equal to present value of income, cash flows, or cost savings actually or hypothetically due to the asset.

Considering all the above, we recognized a few ways to educate our perception about the aforementioned metrics. In order to conclude and to answer the question of how can we enhance our knowledge about all the metrics, in a data driven manner, we summarize the sources or methods to educate ourselves in the following table:

[3] https://www.thebalance.com/how-to-calculate-your-brand-s-value-2295186

METRICS	SOURCES/METHODS TO EDUCATE
Revenue	Direct online channels Business Cases Response Mechanisms
Cost Savings	Business Cases Finance/ Operations Revenue Department
Number of impacted customers	Response Mechanisms Net Promoter Score (NPS)
Loyalty - Repurchase	Retention Metric Historical transactions
Customer Intimacy	Customer Insights User feedback User satisfaction data
User satisfaction effect	Customer Insights User feedback User satisfaction data
Fixed deadline	Management
Dependencies to/from others	Roadmaps Stakeholders/Architects
Security Risks	Security Analysis Security Officers/Architects
Corporate Strategy	Management
Potential penalties/Negative effects	Security Analysis Security Officers/Architects
Value of received information	Customer Insights Development teams Stakeholders/Architects
Business value reduction over time	Business Cases
New Business Opportunities	Social media Stakeholders

Brand Value	Cost based brand valuation Market based brand valuation Income approach brand valuation
PR Value	Secondary data Advisory or consultative groups Unstructured and semi-structured interviews Response Mechanisms Pre-testing
Competitors' Initiation	Benchmarks
Partner Value	Financial fitness Strategic fitness Operational fitness Relationship fitness
B2E Satisfaction	HR Department Employee feedback

Table 13: Means of educating our metrics through data

6.2 Validity issues and limitations

Validity is described as the degree to which a research study measures what it intends to measure. There are two main types of validity, internal and external. Internal validity refers to the validity of the measurement and test itself, whereas external validity refers to the ability to generalize the findings to the target population. Both are very important in analyzing the appropriateness, meaningfulness and usefulness of a research study[4].

In terms of internal validity, there are few aspects that need to be addressed. At first[5], regarding maturation, the calibration of individual opinions as a result of the semi-structured character of the interviews may have an effect on a certain level in our results. Even though there was a certain guideline during the interviews, the discussion would many times escalate or elaborate in specific topics, which could have influenced the perception of the interviewee at that specific point regarding a question.

Moreover, another threat regarding history is that during the transcriptions of the interviews, several parts were missing due to poor recording conditions and/or extra background noises. These parts may not have altered the character of the interview, but they could have contributed to a better perception of a certain answer. Furthermore, regarding experimental mortality, several participants could not elaborate in all three parts of interviews, altering the number of participants through the entire study. Finally, another threat recognized could be

[4] https://psucd8.wordpress.com/2011/11/20/why-is-validity-important-in-research
[5] http://www.indiana.edu/~educy520/sec5982/week_9/520in_ex_validity.pdf

design contamination; in this situation, the participants were aware of the future usage of a metrics framework, thus they evaluated many metrics as important. In case these metrics were not discussed, the results about the proposed framework could have been altered.

As far as external validity is concerned, interaction effect of testing must be taken into account. Pre-testing interacts with the participating population may have caused some effect regarding the results. In addition to that, we should maintain our second thoughts regarding the application of this framework in the industry, as the population, the stakeholders and the deliverables of the organization can be different to many other businesses.

At this point, we have to acknowledge the fact that this research was limited to the Digital Department of KLM – Royal Dutch Airlines. The population of the participants was small, in order to be confident enough to generalize our findings for application in the industry. Nevertheless, the following table depicts the averages of the responses regarding the implementation of the framework, which may provide us with some insights. As a reminder, '1' stands for 'Strongly Agree', '2' stands for 'Agree', '3' stands for 'Neither Agree nor Disagree', '4' stands for 'Disagree', '5' stands for 'Strongly Disagree'.

Statement	Average of responses
I have based my scoring on my knowledge.	**2,57 / 5**
My choice is a gut-feeling.	**2,92 / 5**
I have based my scoring using data.	**3,60 / 5**
I trust that there is related information for this driver.	**2,62 / 5**
I can have access to this information.	**2,28 / 5**

Table 14: Averages of scoring statements, 1= Strongly Agree, 5=Strongly Disagree

Indeed, we notice that the participants were not confident enough to state that they based the majority of their answers on knowledge. Most importantly, we notice that the participants barely based their answers on data, which makes us question the integrity of their scoring regarding the metrics framework. The scoring was based in many cases on the intuition of the participants, which eventually formulated the prioritized lists.

The timeframe used to implement all of three series of interviews within the organization was 3 months, after the conceptualization of the research idea. In that sense, we were not able to conduct more interviews regarding the implementation of the framework, to reinforce its validity. Finally, not all of the initial participants' pool was involved in the final implementation of the proposed metrics framework, making the validation of the PO Matrix even harder.

6.3 Recommendations for future research

It goes without saying that the more implementations of the PO Matrix in for-profit organizations we accomplish, the more confident we can be about its validity. The current sample of the 10 participants within a single organization shows us only the way of such an implementation, and we can only hope that it can be applied even further in different environments and setups.

Moreover, it would be much more helpful if the metrics that are presented to the decision makers are correlated to accessible information; in that sense, the evaluation of each metric would be more realistic and based on data, rather than based on educated guesses.

Finally, and more importantly, in order for the metrics framework to be more effective, it is recommended to apply weights on the categories of the metrics. We have realized that there are several metrics that are taken under consideration more when prioritizing, thus it only makes sense that if a more important weighed metric scores a number equal to another weighed metric, it would accumulate differently in the final calculation.

7. Bibliography - References

Achimugu P., Selarnat A., Ibrahim R. (June 2014), *A systematic literature review of software requirements prioritization research*, Information and Software Technology, Vol.56, Issue 6, pp.568-585, DOI 10.1016/j.infsof.2014.02.001

Ackoff, R. L. (1989), *From data to wisdom,* Journal of Applied Systems Analysis, 16, pp. 3-9

Aksoy, L. 2013. "*How do you Measure what you Can't Define? The Current State of Loyalty Measurement and Management*" Journal of Service Management (24:4), pp. 356-381

Baba, V. V., HakemZadeh, F. 2012. "*Toward a Theory of Evidence Based Decision Making*" Management Decision (50:5), pp. 832-867

Berander P., Khan K., Lehtola L. (2006), *Towards a research framework on Requirements Prioritization,* SERPS'06, Umeå, Sweden.

Brown M. (2013), *Data-Driven Decision Making as a Tool to Improve Software Development Productivity*, Doctor of Philosophy dissertation, Walden University, College of Management and Technology, UMI Number: 3591716

Brynjolfsson E., Hitt L., Kim H., (2011). *Strength in Numbers: How Does Data-Driven Decision making Affect Firm Performance?,* http://dx.doi.org/10.2139/ssrn.1819486

Choppin, J. (2002), *Data use in practice: Examples from the school leve*, American Education Research Association, New Orleans

Confrey J., Makar K. (2005), *Critiquing and improving data use from high stakes tests: Understanding variation and distribution in relation to equity using dynamic statistics software*, pp. 198-226, San Francisco: Jossey-Bass

Cromey, A. (2000). *"Using Student Assessment Data: What Can We Learn from Schools?"* North Central Regional Educational Laboratory, Policy Issues Brief.

Daneva M., Van der Veen E., Amrit C., Ghaisas S., Sikkel K., Kumar R., Ajmeri N., Ramteerthkar U, Wieringa R. (2013), *Agile requirements prioritization in large-scale outsourced system projects: An empirical study*,The journal of Systems and Software, pp. 1333-1353, DOI 10.1016/j.jss.2012.12.046

Davenport T. H., Harris, J. G., and Jacobson, A. L. (2001), *Data to Knowledge to Results: Building an Analytic Capability*, California Management Review (43:2), pp. 117-138

Davenport T. H., Patil D. J. (2012), *Data Scientist: The sexiest job of the 21st century*, Harvard Business Review, October 2012

Denne M., Cleland-Huang J. (2004), *The incremental funding method: Data-Driven Software Development*, IEEE Computer Society IEEE SOFTWARE, 0740-7459/04 IEEE

Drury M., Conboy K., Power K (2012), *Obstacles to decision making in Agile software development teams,* The Journal of Systems and Software 85 (2012) 1239–1254, DOI:10.1016/j.jss.2012.01.058

Dunn K., Airola D., Garrison M. (2013), *Concerns, Knowledge, and Efficacy: An Application of the Teacher Change Model to Data Driven Decision-Making Professional Development*, Vol.4, No.10, 673-682, http://dx.doi.org/10.4236/ce.2013.410096

Garvin, D. A. (2013), "*How Google Sold its Engineers on Management*", Harvard Business Review (91:12), pp. 74-82

Goeken, M. (2011), "*Towards an Evidence-Based Research Approach in Information Systems," in Proceedings of the 32th International Conference on Information Systems"* C. Beath, M. D. Myers, and K. K. Wei (eds.), Shanghai, PRC.

Gupta V., Durg Singh Chauhan, Kamlesh Dutta (2015), *Exploring reprioritization through systematic literature surveys and case studies*, SpringerPlus 4:539, DOI 10.1186/s40064-015-1320-0

Hall D., Song J., Jia L. (2015), "*The Conceptualization of Data-driven Decision Making Capability*", ResearchGate, Conference Paper : Twenty first Americas Conference in Information Systems, Puerto Rico

Haith G., Bowman C. (2014), *Data Driven performance assessment and process management for space situational awareness*, Journal of Aerospace Information Systems, Vol.11, No.3

Hatton S. (2008), *Choosing the "Right prioritization method,* IEEE Computer Society, pp. 517-526, DOI 10.1109/ASWEC.2008.22

Henry J. and Henry S. (1993), *Quantitative assessment of the software maintenance process and requirements volatility,* In Proc. of the ACM Conference on Computer Science, pp. 346–351

Herman, J., & Gribbons, B. (2001), *Lessons learned in using data to support school inquiry and continuous improvement: Final report to the Stuart Foundation* (CSE Technical Report 535), Los Angeles: UCLA Center for the Study of Evaluation

Huergo E. and Jaumandreu J. (2004), *Firms' Age, Process Innovation and Productivity Growth,* International Journal of Industrial Organization, 22(4), pp. 541-59

Ikemoto G., Marsh J (2007). *Cutting Through the "Data-Driven" Mantra: Different Conceptions of Data-Driven Decision Making,* Chapter 5, pp. 105-131

Kaplan, R. S. (2001), *Strategic performance measurement and management in nonprofit organizations. Nonprofit Management & Leadership*, 11, pp. 353–370

Kayes I., Sarker M., Chakareski J. (2013), *Product Backlog Rating: A Case Study On Measuring Test Quality In Scrum, Software Engineering*, https://arxiv.org/abs/1310.2545

Karlsson J., Wohlin C., Regnell B. (1997), *An evaluation of methods for prioritizing software requiremetns,* Information and Software technology 39, pp. 939-947

Kohavi R., Henne R., Sommerfield D. (2007), *Practical Guide to Controlled Experiments on the Web: Listen to Your Customers not to the HiPPO,* KDD: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, p. 959-967, doi : 10.1145/1281192.1281295

Lavalle S., Lesser E., Shockley R., Hopkins M., Kruschwitz N. (2010), *Big Data, Analytics and the Path from Insights to Value*,Vol.52, No.2

Lee G., Xia W. (2010), Toward Agile*: An integrated analysis of quantitative and qualitative field data on software development agility,* Management Information Systems Research Center, http://www.jstor.org/stable/20721416

LeRoux K., & Wright N. S. (2010), *Does performance measurement improve strategic decision making? Findings from a national survey of nonprofit social service agencies*, Nonprofit and Voluntary Sector Quarterly, 39, 571–587.

Light D., Honey M., Mandinach E.B. (2006), *A theoretical framework for Data Driven Decision Making,* EDC Center for Children and Technology, AERA, San Fransisco

Lim C. (2003), *Data for decisions are just a click away*, ENC Focus, 10(1)

Macnamara J. (2015), *PR Metrics: How to Measure Public Relations and Corporate Communication*, https://www.researchgate.net/publication/265317712

Marsh J., Pane J., Hamilton L. (2006). *Making sense of Data-Driven Decision Making in Education,* RAND Corporation OP-170-EDU

Mason S. (2002), *Turning data into knowledge: Lessons from six Milwaukee public schools (WP 2002-3),* Madison: University of Wisconsin, Wisconsin Center for Education Research

Maxwell N., Rotz D., Garcia C., (2015), *Data and Decision Making: Same organization, different perceptions*, Mathematica Policy Research, Working Paper 36

Mazni O., Sharifah A. (2015), *The Impact of Agile Methodology on Software Team's Work Related Well-Being*, International Journal of Software Engineering and Its Applications, Vol. 9, No. 3, pp. 123-136, http://dx.doi.org/10.14257/ijseia.2015.9.3.13

Mingchuo L. (2008), *Structural Equation Modeling for High School Principals' Data-Driven Decision Making: An Analysis of Information Use Environments,* Vol. 44, No. 5, pp. 603-634, DOI: 10.1177/0013161X08321506

Morrel-Samuels P., Francis E., and Shucard S. (2009), *Merged Datasets: An Analytic Tool for Evidence Based Management*, California Management Review (52:1), pp. 120-139

Power D. (2013). *Understanding Data-Driven Decision Support Systems,* Information Systems Management, 25:2, 149-154, DOI: 10.1080/10580530801941124

Rahrovani Y., Kermanshah A., Pinsonneault A. (2014), *Aligning IT for Future Business Value: Conceptualizing IT Project Portfolio Alignment*, The Database for Advances in Information Systems, Volume 45, Number 3

Rami AL-Ta'ani H., Razali R. (2013), *Prioritizing Requirements in Agile Development: A conceptual framework*, Procedia Technology 11, pp. 733-739

Ransbotham S., Kiron D., Kirk Prentice P. (2015), *Minding the Analytics Gap*, MIT Sloan Management Review, pp.63-68, Reprint #56320, http://mitsmr.com/1EkYjyb

Saunders M., Lewis P., Thornhill A. (2012), *Research methods for business students*, 6th ed., ISBN: 978-0-273-75075-8

Serrador P., Pinto J. (2015), *Does Agile work? - A quantitative analysis of agile project success*, International Journal of Project Management 33, pp.1040-1051, https://doi.org/10.1016/j.ijproman.2015.01.006

Setijono D., Dahlgaard J. (2007) *"Customer value as a key performance indicator (KPI) and a key improvement indicator (KII)"*, Measuring Business Excellence, Vol. 11 Issue: 2, pp.44-61, https://doi.org/10.1108/13683040710752733

Schwaber K., Sutherland J. (2016), *The Scrum Guide, The definitive guide to Scrum: The rules of the game*, Scrum.org, Creative Commons

Solomon M. D.(2005), *Ensuring a Successful Data Warehouse Initiative.* Information Systems Management, 22

Stettina C., Horz J. (2015), Agile portfolio management: An empirical perspective on the practice in use, International Journal of Project Management 33, pp. 140 -152, https://doi.org/10.1016/j.ijproman.2014.03.008

Stringfield, S., Wayman, J. C., & Yakimowski-Srebnick, M. E. (2005). *Scaling up data use in classrooms, schools, and districts,* San Fransisco : Jossey-Bass, pp. 133-152

Sulser D. P. (2006) *The relationship between the use of technology for Data Driven Decision Making and student achievement in high school mathematics*, Montana State University, Dissertation of Doctorate of Education, (ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106)

Supovitz, J. A., & Klein, V. (2003), *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*, Report, Philadelphia, PA: Consortium for Policy Research in Education

Provost F., Fawcett T. (2013), *Data Science and its relationship to big data and Data Driven Decision Making,* Mary Ann Liebert Inc., Vol.1, No.1, DOI: 10.1089/big.2013.1508

Tingling, P. M., and Brydon, M. J. (2010), *Is Decision-Based Evidence Making Necessarily Bad?,* MIT Sloan Management Review (51:4), pp. 71-76

Torrecilla-Salinas C.J., Sedeño J., Escalona M.J, Mejías M. (2015), *Estimating, planning and managing Agile Web development projects under a value-based perspective,* Information and Software Technology, Volume 61, pp.124 – 144, https://doi.org/10.1016/j.infsof.2015.01.006

Wang H., Xie M., Goh T.N (2010), *A comparative study of the prioritization matrix method and the analytic hierarchy process technique in quality function deployment,* Total Quality Management, Vol. 9, No.6, pp. 421-430

Wayman, J. C. (2005), *Involving teachers in data-driven decision making: Using computer data systems to support teacher inquiry and reflection,* Journal of Education for Students Placed at Risk, 10(3), 295-308

8. Appendices

List of tables

List of figures