UNIVERSITY OF TURKU

ABSTRACT

	Bachelor's thesis
X	Master's thesis
	Licentiate's thesis
	Doctoral dissertation

Subject	Information systems science	Date	15.06.2020
Author(s)	Antoine Sébé	Number of pages	63+appendices
Title	A decision-making framework for aligning business analytics with business objectives		
Supervisor(s)	Prof. P.H.D. Emiel Caron, Prof. P.H.D. Timo Leino, Prof. P.H.D. Anne-Francoise Rutkowski		

Throughout this thesis, we discuss the impact of Business Analytics on the organizational decision-making process with the objective of designing a framework that provides the organization with extra-knowledge on how to implement and sustain their analytics.

First, we develop the concept of capability using the resource based view and the IT literature to define what is a Business Analytics capabily. We then define the key capabilities that provide the organization with a competitive advantage. Moreover, we investigate the role of governance and alignment as well as the impact of the concepts on the decision making effectiveness. To provide an insight on the adjustment to be made in order to increase the organization Business Analytics performance, we emphasise the role of alignment between Information Technology governance, corporate governance, data governance and Business Analytics governance. Thereafter we create the framework based on academic and empirical research and apply this framework throughout a case study. Based on this case study we provide an academic recommendation to the investigated organization.

This thesis highlight the importance of a the creation of a Business Analyticc governance. Also the reseach provide a framework linking Business Analytics with decision making succesfulness.

Key words	Business Analytics; Decision-making; Governance; Data-driven infrastructure

UNIVERSITY
OF TURKU
Turku School of
Economics

A DECISION-MAKING FRAMEWORK FOR ALIGNING BUSINESS ANALYTICS WITH BUSINESS OBJECTIVES

Master's Thesis
in Department of Management and Entrepreneurship: Information Systems Science

Author:
Antoine Sebe

Supervisor:
Prof. Ph.D. Emiel Caron
P.H.D. Timo Leino, Prof.

15.06.2020
Turku

AKNOWLEDGEMENTS

In order to complete the International Master of Information Technology based in Tilburg University, Turku University and the IAE Aix-Marseille I was required to produce a thesis based on the topic studied throughout the program. Moreover in addition of this thesis I joined Amazon as part of an internship. This master thesis embodies a research done at Amazon.

The subject of this thesis is related to Business Analytics. I chose this topic because the class of Business Intelligence and Emerging Trends followed at Tilburg University has attracted my curiosity and led me to seeking more information about the topic. Moreover, my job at Amazon closely related to business analytics made this topic a perfect fit for this academic work.

I must say that writing a thesis and performing a full-time internship has been rather stressful and sometimes overwhelming. But I feel that I have been constantly supported throughout this journey. Therefore, I would like to thank the people that helped me carry on when my motivation what at its lowest. First I would like to thank my manager at Amazon that showed me the way toward a potential future carrier at Amazon. Then I would like to thank my girlfriend that supported me and never let me down during this difficult moments. And finally I would like to thank my thesis supervisor that challenged me but also motivated in providing the best output, I trust that the feed-back given along my writing have consistently improved the quality of this thesis.

Antoine Sebe, June 2020

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background

During the past two decades, the importance of Business Analytics (BA) has increased exponentially and has become a key factor of success for organizations. Larson (2016) emphasizes that analytics is now a core enabler for the entity to make a better decision and to improve organizational efficiency. Indeed, the role of analytics has evolved from simple statistical analytics for operational decisions to strategic planning, customer relationship management and process development (Negash, 2008).

In the meantime, as the role of BA was changing the complexity of analytics techniques followed the same trend and shifted from basic tools like dashboards to much more sophisticated techniques such as machine learning (Olszak, 2012). It is important to notice that as the discipline grew the size of the data to be processed followed the same evolution which embodies the first organizational challenge. Indeed, technology allows the organization to collect a wider range of data from various sources, therefore, the requirement for data processing capacity consistently increases. However, processing big data is not the only challenge that organizations must cope with to reach performance using BA.

Even if the analytical part of BA is widely acknowledged, the steps needed to efficiently provide decision-makers with the relevant data are still not clearly defined (Labrinidis, 2012). Therefore, when the objective of the organization using BA is not well established, the use of the discipline becomes risky and the benefits expected from the technology are often not achieved (Heinrich, 2003). For instance, in some cases, organizations collect an important amount of data but do not have clear governance that guides the use of this data to create business value. This is why an organization-wide strategy and governance is required to integrate BA into the business process of the organization in order to allow decision-makers to fully grasp the benefits of the discipline.

Therefore, the three key components necessary to overcome the challenges that keep the organization away from successfully enhancing decision making using BA are technology, data processing and governance. The focus of this thesis is to investigate the influence between these dimensions and to provide an insight into how the organization can best leverage its data to create value.

To do so, we investigate the worldwide organization Amazon recognized both for heavily monitoring its internal business processes using BA and for its ability to create an accurate customer profile based on online shared data. The profile of the organization will provide a broad scope of analysis for this research as the organization covers a wide range of activity powered by BA.

1.2 Research questions

Based on the introduction the main research question is **"How to design a decision-making framework that links Business Analytics with the objectives of the organization?"**

To answer this main research question, five sub-questions have been developed:

(1)What is Business Analytics and what are the key concepts and tools related?

(2)What are the key capabilities leading to Business Analytics successful implementation?

(3)How to identify Business Analytics strategy misalignments? Is Business Analytics alignment a capability?

(4)Is Business Analytics governance a capability?

(5)How to validate the decision-making framework?

(6)How to apply a decision-making framework to Amazon?

1.3 Research relevance

1.3.1 Business relevance

The first objective of this research is to design a framework that provides an overview on how to successfully improve decision-making using BA. There are currently only few frameworks that link the use of BA to decision-making effectiveness. Therefore this framework will provide the organization with a first insight on how to harness BA in order to develop enhance decision-making. Also, this research aims at increasing organizational awareness concerning BA governance/alignment inhibitors and good practices. Practicaly speaking the created framework aims at providing managers and top manager with a clear insight on what are all the dimensions and components to be taken into consideration in order to reach BA performance. Moreover mapping the inter-connections of the components, the framework enhances the organization understanding of the BA's dimensions interdependencies.

1.3.2 Scientific relevance

Even if the concepts of IT governance, alignment and Business Analytics have been individually heavily studied across the past two decades, there are very few academic literature concerning BA alignment and governance. Moreover, there is no literature relate to BA misalignment. This research use an existing succeful ERP framework in order to provide a first insight on what are the BA misalignment present across the organization, therefore filling the gap in the BA misalignment literature.

1.4 Research methodology

1.4.1 Design research model

Design-science research aims at improving the organizational performance by creating artefacts that "*define ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of information technologies can be effectively and efficiently accomplished*" (Denning, 1997; Tsichritzis, 1998).
The objective of this design science research (Appendix 1) is to create a framework that supports the organization use of BA. The research starts by identifying the BA needs of the organization (Figure 1).

Thereafter, the artefact is designed by integrating both the needs of the organization and academic literature. After the first design of the framework, BA experts are interviewed and and given a survey. During the interviews, the experts were given the framework and asked for their feedback. Based on the expert feedbacks, the framework was redesign, improved and experts validated the framework. Furthermore, the validated framework increased experts knowledge but did not improve the organization's efficiency. To provide the organization with an insight on its current state in terms of BA, the framework was applied to the organization and a gap analysis has been developed. Using the findings gathered throughout the gap analysis, recommendations based on academic literature have been developed to help the organization achieve its desired state.

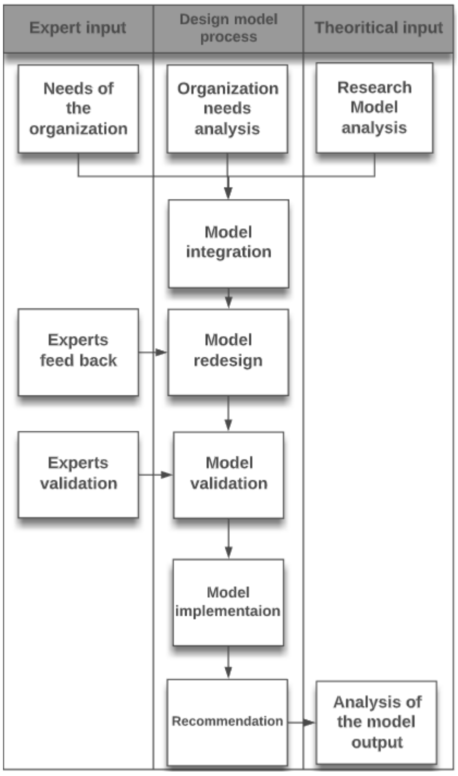


Figure 1 Framework design process

1.4.2 Expert selection

To optimize the efficiency of the to-be-created framework, experts have been selected based on their role and experience. The panel of experts was composed of managers covering the entire scope of BA. The first group was composed of two senior managers working closely with BA that have been in the company for at least three years. The second group was constituted of three Senior Business Analysts that have been working within the company for at least two years. The third group detains two Data Scientists that stayed in the organization for at least one year.

All interviewed experts' identities remained anonymous and are mentioned as Business Analyst 1, 2, 3, Data Scientist 1, 2,3 and Senior Manager 1 and 2.

Experts were randomly selected and the only criterion taken into consideration during the selection is their role within the organization.

1.4.3 Experts interviews

As priory mentioned, the first phase of the research consisted of conducting semi-structured interviews to improve the new framework.

As part of this semi-structured phase, experts were given the frameworks. Thereafter, they could provide their opinion on the framework by pointing out the missing parts.

In a second time, the experts were given a survey. The purpose of this second part is to assess the BA aspects of the framework based on:

(1) The importance of the component for the organization

(2) The performance of the component within the organization

The components have been graded on a scale from one to ten; one representing a component that either has no importance for the organization or was not present; and ten representing a component being fully performed or of critical importance for the organization.

1.4.4 Research reliability and validity

Aming at increasing the research validity, we use triangulation by combining different methods to collect data and triangulate the results.

The first method used to collect data was a survey that was aplied throughout interviewes. Indeed, solely using interviews would only have provided the experts' opinion which is important but limited in terms of data validity. Moreover, only using a survey could have led to biases in answering the survey.

Therefore, to increase the research validity, experts were asked to both grade the components but also to justify their grade by providing facts and numbers.

Moreover, the data concerning the framework is collected throughout academic research which implies that both survey and interview are based on former validated research.

The sample of selected experts covers most of the scope of the topic: Senior managers working in BA-driven environment, Business Analyst, Data Scientist. All the experts have been dealing with Business Analytics during both their university studies and throughout multiple projects which consistently diminished a potential bias linked to a lack of experience.

Those different competencies and the experience of the respondents both within and outside the company increased the external validity of the research.

Concerning the internal validity, experts have been randomly picked based on their function within the organization. Moreover, each expert has been given the same survey and was only asked to justify its choices using the same pre-established questions which increases the internal validity of the study. Also, interviews results are included in appendix X as an audit trail which improves the internal validity.

1.5 Research scope

Even if this research is based on all the dimensions of BA capabilities, the recommendations made to Amazon will only be based on the three dimensions: data-driven infrastructure (i), data-driven decision-making (ii) , data processing (iii) and BA technologies (iv).

Because the framework developed cannot be extensively tested, the research is solely exploratory. This is due to the time constraint of approximately four months given by the universities of Tilburg, Turku and Aix-Marseille.

Also, the foundation for the recommendations are only based on academic evidences and are not supported by clear instructions on how to implement them. Furthermore, the only organization investigated during this case study is Amazon. Finally, this research is based on an overall view of the BA and discuss the different capabilities related to the concept.

1.6 Research outline

Following the herein part of this thesis, the second chapter provides a definition of BA but also of BA capability. The main objective of the chapter is to provide a solid insight into what are the main BA capabilities. The definition of BA capability is based on the concept of IS capability.

Chapter 3 explains all the core concepts and tools related to BA. The tools described throughout this chapter are picked based on the organization needs for BA technologies. One of the most important concepts of this chapter is data quality that is highly related to the concept of BA capabilities. Thereafter, chapter 4 points out the need for BA alignment based on the research for IS alignment. Also, the research on ERP system is used to provide a first insight on what are the BA misalignment criteria and consequences.

Finally, the fifth chapter provides a definition of IT, data and BA Governance. By integrating chapters 2, 3 and 4, the research model is created and presented in this fifth chapter. The research model highlights the relationship between all the dimensions defined in the previous chapters. Moreover, chapter 6 consists of defining the decision-making framework by explaining its components. Also, using the feedback provided by the expert the framework is validated. Chapter 7 consists of actually applying the framework to the organization by generating a gap analysis between the current and desired state for each component of the framework. Using the outcome of the gap analysis, a set of recommendation is provided.Then, chapter 8 provides an answer to the research question based on the six answers given to the sub-research questions. Finally, the ninth chapter emphasizes the limitations, further implications of the result and ends with the recommendation for further research.

2 BUSINESS ANALYTICS OVERVIEW

Based on a literature review of the various definition for the concept, this chapter defines what is BA. In this chapter, we first define the key concepts related to BA starting from the broadest concept: data quality and then narrowing the scope of the definition. In a second time, the tools that are assessed throughout the case study in chapter 7 are defined in section 2.6.

2.1 Defining Business Analytics

Even if the concept of BA is composed of many dimensions that will be discussed later in this research, this part intends to define the core of BA. As BA is growing, the different definitions of the concept are also consistently evolving.

According to Holsapple, C. (2014) the definition of BA can be divided into six classes, each containing a peculiar emphasis (Table 1). However, the keywords "fact-based" and "decision making" are mentioned in almost all the definitions. Therefore, even though the interpretation and the rationale behind BA definition are somehow different, it can be agreed that the foundation of BA is built on a "fact-based decision making". This is why throughout this research we will rely on this definition of BA that defines the discipline as the process of exploiting data not only to improve decision making but as a base for any operational/tactical/strategical decisions.

Also we that BA comprise three core clusters of analysis (Manyika, 2011):

- Predictive analytics: this type of analysis uses advanced analytics tools in order to predict future unkown events.
- Prescriptive analytics: this type of analytics uses data related past and current performance in order to create a scenario to propose a strategy.
- Descriptive analytics: This analysis use historical data in order to explain the changes that impacted the organization.

Table 1. Holsapple, C. (2014)

Class	Definition of BA	Rational for BA
A Movement	*"management philosophy, through which insights can be gained and decision making improved based on a rich set of data"*	*"improve the overall process or decisions associated with process roles"*
A Collection of Practices & Technologies	*"a group of tools that are used in combination with one another to gain information, analyze that information, and predict outcomes of the problem solutions"*	*"allow for informed decision making"*
A Transformation Process	*"the scientific process of transforming data into insight for making better decisions"*	*"decisions and insights obtained from analytics are implemented through changes within enterprise systems"*
A Capability Set	*"extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions"*	*"help your managers and employees make better decisions, and help your organization perform better"*
A Decisional Paradigm	*. "accessing, aggregating, and analyzing large amounts of data from diverse sources to understand historical performance or behaviour, or to predict or manage—outcomes"*	*"rapid implementation of new ideas, products, and services, which result in greater profits and shareholder value"*
Specific Activities	*"[the]part of decision management" that involves "logical analysis based on data to make better decisions"*	*"create more value for customers and more profit for companies"*

2.2 Data quality

We highlight that data quality represent a critical foundation for BA. Indeed, in order to harness the power of the discipline, qualitative data is a basic requirement and therefore is one of the most important concept of BA.

2.2.1 Definition of data quality

From a business perspective, Redman (2013) defines qualitative data as the data that "*fit for its intended uses in operation, decision making, and planning*". From a more general perspective, Wang & Strong (1996) emphasize that the data quality depends on the needs of the data users. Therefore, if data quality has multiple dimensions, the organization must identify the dimensions that are most critical for both maintaining the business and producing innovation. In this thesis, we argue that an organization must identify the data dimension corresponding to its needs in order to develop the capability of capturing, integrating, analyzing and displaying data.

2.2.2 Characteristics and dimensions of data quality

Even though the early literature related to data quality named fifteen different dimensions of data quality (Strong, Lee & Wang, 1997), the development of big data made those dimensions irrelevant (Cai&Zhu, 2015). Therefore, Cai And Zhu (2015) propose a new division of the data quality criteria that is composed of five dimensions. Each dimension is divided into criteria that are explained below (Table 2).

Table 2. Data quality dimensions (Cai, X. & Zhu, X., 2015)

Dimension	Elements	Definition
Availability	Accessibility	*”Accessibility refers to the difficulty level for users to obtain data.”*
	Timeliness	*”Timeliness is defined as the time delay from data generation and acquisition to utilization (McGivray, 2010).”*
	Authorization	*”Authorization refers to whether an individual or organization has the right to use the data.”*
Usability	Credibility	*”Credibility is used to evaluate non-numeric data. It refers to the objective and subjective components of the believability of a source or message. The credibility of data has three key factors: reliability of data sources, data normalization, and the time when the data are produced.”*
	MetaData	*”With the increase of data sources and data types, because data consumers distort the meaning of common terminology and concepts of data, using data may bring risks. Therefore, data producers need to provide metadata describing different aspects of the datasets.”*
	Definition/ Documentation	*”Definition/document consists of data specification, which includes data-name, definition, ranges of valid values, standard formats, business rules, etc.”*
Reliability	Accuracy	*”To ascertain the accuracy of a given data value, it is compared to a known reference value.”*
	Integrity	*”In a database, data with “integrity” are said to have a complete structure. Data values are standardized according to a data model and/or data type. All characteristics of the data must be correct – including business rules, relations, dates, definitions, etc.”*
	Consistency	*”Data consistency refers to whether the logical relationship between correlated data is correct and complete.”*
	Completeness	*”If a datum has multiple components, we can describe the quality with completeness. Completeness means that the values of all components of a single datum are valid.”*
	Auditability	*”From the perspective of audit application, the data life cycle includes three phases: data generation, data collection, and data use (Wang & Zhu, 2007). But here auditability means that auditors can fairly evaluate data accuracy and integrity within rational time and manpower limits during the data use phase”*
Relevance	Fitness	*”Fitness has two-level requirements: 1) the amount of accessed data used by users and 2) the degree to which the data produced matches users’ needs.”*
Presentation Quality	Readability	*”Readability is defined as the ability of data content to be correctly explained according to known or well defined terms, attributes, units, codes, abbreviations, or other information.”*
	Structure	*”More than 80% of all data is unstructured, therefore, structure refers to the level of difficulty in transforming semi-structured or unstructured data to structured data through technology.”*

The dimensions mentioned above are critical for small and medium data. Nevertheless, while dealing with big data Katal, Wazid and Godar (2013) identify six key characteristics of big data (Table 3).

Table 3. Big data characteristics (Katal, X. Wazid, X., & Gobar, X., 2013)

Volume	The huge amount of data constantly generated, integrated and analyzed.
Velocity	The high speed of data coming from a different source with a different format.
Variability	The variability of the data defines the variation in the data flow. This creates a challenge in maintaining the data loads.
Complexity	The complexity of big data comes from the difficulty arising from the velocity, the variety and above all the huge volume of data.
Value	Big data create the possibility to capture a huge number of valuable data, therefore, improving the organization decision making process.
Variety	Due to the increasing number of data format sources (images, social media, sensors etc…). The variety of unstructured, raw, structured and refined data create a real challenge for a data miner to rip the benefices of big data.

2.3 Big data

Big data defines large and complex set of data with diverse sources of unconnected sources (Wu, 2013).The concept of big data is embodied by three key attribues: (1) the size of the data; (2) the complexity of the data in term of structure, permutation, uniformity and behaviour ;(3) and the advanced technologies and tools necessary to process sizable and complex dataset.

For databases to be able to deal with big data, one of the most critical elements is the database scalability (Wu, 2000). In its research, he proposes a three-tier big data processing framework that describes the different challenges at each for each tier:

- Tier I: data accessing and computing procedures
- Tier II: data privacy and domain knowledge
- Tier III: big data mining algorithms

Those three concepts are defined in Appendix 2.

2.4 Data mining

Data mining can also be called "*knowledge discovery in databases*" and design the process of finding new patterns an relationships in databases that will support decision-making (Bose, 2001). Accroding to Liao (2003) represents the interdisciplinary field that "Combine disciplines such as statistics, database management, machine learning, computer science and artificial intelligence."

2.5 Business analytics tools

In this section a comprehensive approach of BA tools. Note that a more elaborated description can be found in the appendix.

2.5.1 Machine learning application in businesses

Machine learning is a sub field of data mining and is used to analyze data and discover useful patterns. Machine learning is the study that aims at automating the knowledge discovery process (Bose, 2001). This method permits to eliminate time-consuming data by engineering systems that develop knowledge. One of the common machine learning processes involves identifying a recurrent pattern in a training data set and forecast the behaviour of other similar data sets. The definition of the three machine learning techniques used at Amazon can be found in Appendix 4. We note that machine learning is part of the category predictive and prescriptive analytics.

2.5.2 Process mining

The process mining field finds its foundations in information technologies like Enterprise Resource Systems, Customer Relationship Management systems, Supply Chain Management systems and other B2B systems. Those process management systems use business event called “event logs”. An event log is composed of different features, in the first place each event corresponds to an activity. Moreover, each event refers to a case and may have one or several performers. Finally, each event detains ordered timestamps. Process mining methods use event logs for process discovery, verify compliance, identify flaws and differentiate the process paths (Van Der Aalst, 2016). From an organizational perspective, the data related to the performer permits the company to increase its understanding concerning the different involvement and relationship of the employees that make use of systems (who uses the process). Further objectives of process mining are defined in Appendix 3.

2.5.3 Web Analytics

During the past twenty years, the increasing use of web-based technologies has completely changed the way companies gather, store and use their data. In this context, Web

Analytics has been defined as "*the measurement, collection, analysis and reporting of Internet data for understanding and optimizing Web usage*" (Lakshmi S. Iyer, 2011). Even if Web Analytics' first application aimed at measuring the internet traffic and optimizing online sites, the current scope of the practice is much more oriented toward improving strategical and tactical decision-making (marketing research, customer segmentation etc.) (Appendix 5). We note that web analytics is mostly part of descriptive analytics.

2.5.4 Dashboard

Due to the increasing amount of information, data overload has been identified as one of the current negative trends in many companies consistently using information technologies. To manage this overload, dashboards are a very efficient tool. Dashboards are regularly used as a management tool to measure the performance and the advancement of a project. This tool can be used to cover a wide area of activities from monitoring the performance of the organization strategy to tracking the capability of a unit to reach its service level agreement. Even if dashboards are currently widely used, they can be hard to develop and often don't permit to achieve their monitoring objectives (Kawamoto, 2007). We note that web analytics is part of descriptive analytics.

2.5.5 Social media analytics

Social media analytics is a quickly evolving capability that permits the organization to capture and analyze a large quantity of online data to identify behaviours and opinion trends.

The objective of the discipline consists in integrating and developing analytical methods aiming at mapping the social media trends (Stieglitz, 2014). Social media analytics sustain a wide range of other disciplines like IS and BA. Holsapple (2014) explains that "*business Social Media Analytics refers to all activities related to gathering relevant social media data, analyzing the gathered data, and disseminating findings as appropriate to support business activities such as intelligence gathering, insight generation, sense-making, problem recognition/opportunity detection, problem-solution/opportunity exploitation, and/or decision making undertaken in response to sensed business needs*". Social media analytics trends and key factor of success are analysed in Appendix 6. We note that web analytics is mostly part of descriptive analytics.

3 BUSINESS ANALYTICS MANAGERIAL CAPABILITIES

Along this chapter, the definition of an IT capability is used to build the definition of BA capability. Then, the dimensions of BA capability are unveiled. Within this chapter we discuss the different BA capability that are critical for reaching decision-making effectiveness from a managerial perspective. We also emphasize that BA managerial capabilities depends on a higher level of capability discussed in chatper 4&5.

3.1 Information Technology capability

The foundation of IT capability is based on the resource-based view (Appendix 7) that emphasizes that an organization create a competitive advantage by combining resources that are *"non-substitutable, scarce, difficult to imitate and economically valuable"* (Barney, 1991). Furthermore, Bharadwaj (2000) suggest that IT capability is the "*firms' ability to mobilize and deploy IT-based resources in combination or co-present with other resources and capabilities"*.

The study on IT capability has evolved from a single dimension (technological capability) to three dimensions including the technological dimension (i), the human dimension (ii) and the organizational dimension (iii) (Kim, 2011). The technological dimension defines the structure and the settings of all the technological components of an organization which comprises software, hardware and applications. Then, the human dimension refers to the knowledge and skills of the IT employees as well as the organization's ability to leverage IT as a competitive advantage. Finally, the organizational dimension represents the IT-business alignment that provides the organization with a competitive advantage (Rockmann, 2014).

Therefore, IT capabilities represent the organization's ability to implement and use its IT systems efficiently.

3.2 Definition Business Analytics capability

Using the former definition of IT capability, we can define what a BA capability is. BA as a capability refers to an operational capability and can be conceptualized as a set of processes that add a second layer of capability to the previously described IT capability (Isik, 2013). This second layer of capability consists of four dimensions: (i) the data-driven infrastructure dimensions that aims at creating a (ii) data-driven environment, the data processing dimension (iii) that refers to the organization's ability to integrate data and intelligence, and the decision making capability dimension (iv) which represents the BA capability to improve customer and business understanding (Ramakrishnan, 2018). We note that the dimension data-driven infrastructure influences both the data processing dimension as well as the decision-making capability dimension. This definition of BA capability is based on an information processing view that fosters the relevance of matching the information processing requirement with the information processing capabilities.

Also the process view emphasizes that the "*organisations should design its structure or business processes to enable decision-makers to process a great amount of data, thereby to inform decision-making, reduce costs, and improve organisational performance*" (Premkumar, 2005).

3.3 Business Analytics capability dimensions

In this section we define the different managerial dimensions of BA as well as all the components relative to these dimensions(Table 4). We note that the BA managerial capabilities depend on the organizational ability to align BA with the need of the business.

Table 4. Business capabilities components (adapted from Cao, 2015)

Business Analytics capabilities		
Dimensions	Components	Sources
Data-driven infrastructure	1. The organizational structure is adapted to BA 2. The entity's structure consistently supports BA. 3. BA is fully integrated within all the core processes of the company. 4. The company's policies and rules guide the BA.	(Davenport, 2001) (Kiron, Prentice, 2012) (Kiron, 2011) (Davenporty, 2013)
Data processing capabilities	5. The entity is capable to efficiently capture data. 6. The entity is capable to efficiently integrate data. 7. The entity is capable to efficiently analyse data.	(Davenport, 2001) (Lavalle, 2011) (Kiron, 2011) (Kiron, 2012)
Data-driven decision making	8. The entity is capable of relying almost exclusively on BA for decision making. 9. The entity is capable of using BA for innovating. 10. The entity uses BA to challenge current practices.	(Kiron, 2012) (Davenporty, 2013) (Kiron, 2011)
Decision-making effectiveness	11. The entity is capable of quickly responding to changes using BA. 12. The entity makes a real-time decision using BA. 13. The entity is capable of understanding the client requirement using BA.	(Lavalle, Lesser, 2011) (Kiron, 2012) (Kiron, 2011)
Business Analytics technologies	14. The entity detains all the required human capabilities. 15. The entity detains all the required technological capabilities.	(Chen, 2012) (Davenport, Harris, 2001) (Lavalle, 2011)

3.4 Data-driven infrastructure dimension

The data-driven infrastructure refers to three objectives technical, structural and cultural. The first objective refers to the organizations' technical readiness which includes analytics, collaboration, innovation as well as security and data privacy. Then, the structural objective relate to the organization ability to organize the cultural and technical dimension with the purpose of reaching business process implementation and innovation. Finally, the cultural objective defines the organization capability to manage data, processes and knowledge to create a dynamic communication between stakeholders (Ramakrishnan, 2018). Lavalle (2011) suggests that to produce a data-driven decision-making process the business strategy must be tightly coupled with the analytic staretegy and easily understandable from a user perspective. Analytics should be integrated into the core business processes of the organization and allow decisions to be taken in a timely fashion. Furthermore, it is critical to developing a clear entity structure that closely integrates BA into the company processes as well as clear rules and policies that should guide BA (Cao, 2015).

Based on the previously mentioned literature the four components related to the data-driven infrastructure capability are developed: the organizational structure is adapted to BA (1), the entity's structure consistently support BA (2), BA is fully integrated within all the core processes of the company (3), the company's policies and rules guides BA (4).

We also note that to enforce these components, a BA governance is required to align the strategy of the organization with the BA needs. The concept of alignment will be discussed in chapter 4 and the concept of governance in chapter 5.

3.5 Data processing capability

The data processing component is the ability the organization has to lead its projects rolling out its data analytics capability. This capability becomes a competitive advantage for the organization when it is managed in such a way that the organization can efficiently implement a complex project in a changing environment leveraging data.

Data processing consists in a range of tasks that include the design and maintenance of the organizational network, the acquisition of technical expertise, the transformation of data into useful information and the continuous integration of technical and human intelligence into the processes of the organization (Petrini, 2009).

Moreover, in order to further develop this capability, it has been shown that one of the key criteria is the organization ability to align its data processing resources and the needs of the business (Cao, 2015). The data processing capability of an organization is defined as its ability to capture, integrate and analyse data to enhance decision-making process and outcome.

Finally, a critical factor for the successful creation of a data processing capability is the organization's ability to maintain and rely on qualitative data. The components reflecting this capability are:

- the entity is capable to efficiently capture data (5);
- the entity is capable to efficiently integrate data(6);
- the entity is capable to efficiently analyse data (7).

It is important to notice that the outcome of this capability permit the enterprise to produce qualitative data that will allow the company to produce an effective decision-making, the concept of data quality is discussed in section 3.1.

3.6 Data driven decision making capability & effectiveness

According to Isik (2013), BA is critical to support decision-making within any organization. The data-driven decision-making dimension refers to the company's ability to leverage its data using BA tools and the data-driven infrastructure to produce decision making effectiveness. Data-driven decision-making can be described as the organization's ability to challenge the status quo and monitor the business based on insight-driven decisions (Cao, 2015). This definition suggests that the organization should heavily rely on BA to take decisions.

Moreover, the company must be able to leverage its tools to extract useful information that would permit to create new business processes and to innovate in its market.

Also, the company's ability to base its decision on data is strongly correlated with its ability to integrate BA into its business process which enables the identification of sub-efficient practices. Therefore decision making effectiveness strongly depends on the organization's BA capability to process data and to harness this data capability by aligning BA with the entity's objectives (Cao, 2015). The components reflecting these capabilities are the following: the entity is capable of relying almost exclusively on BA for decision making (8), the entity is capable of using BA for innovating (9), the entity uses BA to challenge current practices (10), the entity is capable of quickly responding to changes using BA (11), the entity makes real-time decision using BA (12) and the entity is capable of understanding the client requirement using BA (13).

3.7 Business Analytics technology capability

The variety and sophistication of BA technology can provide a competitive advantage to the organization. Nevertheless, it is argued that the organization's ability to leverage BA tools and technologies can be strongly diminished if the organization doesn't support BA with an adequate structure (Chen, 2012). Moreover, to promote a collective intelligence the organization should foster sharing and foster cross-functional teams. We argue that there is a common set of tools that fit the needs of every organization.

Finally, the technology capability is based both on (14) human and (15) technological capabilities. Therefore, the major component of the following capability is the entity's' ability to gather all the BA capabilities in terms of human and technical resources.

Also, it is important to notice that BA technology must be closely aligned with the need of the organization (Chen, 2012).

4 ALIGNMENT

Firstly, we define the concept of IT/Business alignment and misalignment. Then, the concept of ERP misalignment is discussed. Based on the definition of IT alignment/misalignment we define what BA alignment is. Moreover, we leverage the misalignment component suggested throughout the ERP literature to propose a set of BA misalignment.

4.1 IT and Business Alignment

4.1.1 Information Technologies alignment definition

In the Contingency view, “alignment” is defined as “*the degree to which the needs, demands, goals, objectives, and/or structures of one component are consistent with the needs, demands, goals, objectives, and/or structures of another component*” (Nadler, 1986). More precisely, El-Tebany (2014) defines the alignment between business and IT as “*the extent to which information technologies support and have a positive relationship with the organization’s objectives and strategies as defined in the business plan in an appropriate and timely way*”. As the business environment is growing in term of complexity, the shift between organization objectives and its IT has become a primary topic of research. Therefore, across the past two decades, many frameworks aiming at aligning the IT with the goals of the entity have been created.

One of the most commonly accepted model for IT alignment is the Strategic Alignment model proposed by Venkatraman (1993). This model is based on the assumption that an effective management of IT is based on four alignment perspectives. (1) The first perspective emphasizes that the Business Stateggy is structured in such a way that it drives both the organizational design and the IT infrastructure. Within this set up, the top managers are responsible for designing most of the strategy, and IT managers to implement it .(2) The *“technology transformation”* refers to implementing a business strategy throughout the IT strategy. We note that these two perspectives are driven by the business side. (3) The *“competitive potential”* perspective refers to the exploitation of emerging technologies capabilities, therefore, the IT strategy is a critical attribute influencing the business strategy as well as the organization infrastructure. (4) The “service level” perspective emphasize the creation of a world-class IT service and the IT strategy drives both the IT and organizational infrastructure. In this case the business strategy can be seen as secondary.

4.1.2 Information Technologies misalignment definition

Even if the alignment between Business and IT has been extensively researched, the organization ability to reach this state has been sub-efficient (Lederer, 2009). Although former literature accurately provides frameworks to reach alignment, only very few researches focus on identifying misalignment that impedes the organization to reach its IT goals. According to Luftman (2003), a misalignment is *"a set of symptoms or factors that organization might experience indicating that the structure is not optimized"*. Even if this definition is very general it unveils two characteristics of a misalignment: (1) a misalignment can emerge as a symptom; (2) a misalignment prevent the organization from reach its full performance. The presence of a symptom within the organization can indicate that the alignment of business-IT strategy is not fully reached. This view shifts from the previous alignment research, as, instead of trying to provide a method to align business and IT, it intends to identify the misalignment that prevents the business to reach this alignment. Thus, El-Telbany (2014) defines Business-IT misalignment as "*the continuous efforts, involving management and information technologies, of consciously and coherently detecting and testing for the interrelation of all components of the business-IT relationship; where a change in one would instantly influence the other, contributing to the organization's performance over time"*. To enhance the capabilities of an organization to reach alignment, this research emphasizes the need for a mechanism that aims at identifying and resolving the misalignments. We argue that to reach alignment, both alignment and misalignment perspectives must be incorporated into the company IT strategy.

4.1.3 The five construct of IT alignment

To identify what business-IT alignment looks like, El-Telbany (2014) suggests five main constructs of IS alignment being: *"Business-IT Relationship, IT project, Business-IT communication, Business-IT engagement"*. The Business-IT Relationship embodies the link between business and IT. This construct is closely related to the development of an IT vision broadcasted efficiently throughout the organization as well as the availability of IT resources. The IT project construct represents the creation phase of an IT project. It refers to the extent to which an IT project is in accordance with the business objectives but also if the human, technical and financial resources are aligned with the goal of the IT project. The business-IT communication construct represents the communication channel established between IT and business. This construct regards the frequency of the communication between IT and business and the interaction process (one-way interactions versus two ways interactions, formal or informal). Finally, the Business-IT engagement is the agreed deliverable services between business and IT. The last construct regards *Business-IT Misalignment.* It defines the availability of a procedure to identify, investigate and solve misalignment.

4.1.4 Misalignment between the Information Technologies/Business a BITAM approach

Sousa (2008) provides a misalignment conceptualization at the structural level: the Business and IT Alignment Model (BITAM). This model is based on three levels and around the concepts of Business Model, Business Architecture and IT Architecture. The BITAM approach refers to three stages of maturity that relate to the organization ability to cope with misalignments:

- Detection
- Correction
- Prevention

According to the BITAM analysis, each level of maturity is based on the former one. For instance, to correct misalignment, the organization must be able to detect it. The diagnosis defines the procedures of detecting a misalignment using its signs and symptoms as a result of a questionnaire or a test. Once the misalignment has been detected the entity proceeds to the correction using therapy. The process of therapy consists of solving the etiology by understanding the symptoms in order to re-align the business with IT. For instance, one example of misalignment therapy could be the definition and the communication of the entity's objectives and strategy. Thereafter, the prevention phase is the final objective of the organization. The process of ensuring IT/business alignment by preventing misalignment is called prophylaxis and is directly related to the organizational capability of detecting and correcting misalignment in an efficient and timely manner. Organizations able to efficiently forecast future misalignment and adjust strategy accordingly are the most likely to succeed in aligning IT with the goal of the organization (Appendix 8. Definition misalignment medical semantic).

4.2 Misalignment between ERP and the goal of the organization

4.2.1 The role of an ERP System

ERP represents the most complete and efficient enterprise system. This system enables the company to integrate the different sectors and department that drive the organization. This software application permits the organization to pilot its processes with an integrated system. Nevertheless, the lack of capability of the company to efficiently harness the benefits of this solution can result in a very costly failure (Van Groenendaal, 2015).

4.2.2 ERP misalignment

According to Rosemann et al. (2004), a misalignment in ERP systems embodies the difference between the needs of the organization and the solution capabilities. Thus, an ERP misalignment can have two etiologies, either the feature requested by the organization is missing from the package or the solution imposes an inefficient business process. A misalignment refers to a sub-optimal fit between the IT strategy embed in the ERP system and the actual strategy of the organization. In other words, a misfit, or misalignment results from the gap between the organization's needs and the ERP selected features or price (Yen et al., 2011).

4.2.3 Linking Enterprise ERP literature with Business Analytics

Information technologies like Enterprise resource planning (ERP) have provided the organization with the possibility to develop their capability to capture, process and analyze data (Appelbaum, 2017). Indeed, such systems allow the company to extract data and meta-data from both external and internal sources. These systems provide support to drive performance measurement and to increase evidence-based decision-making. Because the littarature on BA misalignments is nihil, in the below section we consider that the purpose and the features of ERPs are convergent with BA and therefore we base our analysis of BA misalignment on the former ERP literature. Moreover, the misalignments proposed in the ERP literature are based on the software application perspective that is based on three categories: process, data and output.

4.3 Business Analytics alignment

Event if the COBIT (*Control Objectives for Information Related Technology*) provides a first insight on good practice to aligned analytics with the goal of the organization, there is currenly very few literature that aims at identifying BA misalignment, therefore, the identification of BA misalignment is based on the combined literature produced in section 4.1, 4.2, 4.3 as well as section 2.6.

4.3.1 Definition Business Analytics Alignment

Before the application of the concept of alignment to BA, the contingency view has been heavily used to investigate the synergy between IT and the organisational dimensions. This concept emphasizes that when business and IT dimension are embedded they will create IT capabilities. These IT capabilities will provide the organization with the ability to leverage its technologies and to create a competitive advantage. Also, Cao (2015) shows that the alignment of BA and the organization infrastructure has a positive impact on the companies' decision making capability (Cao, 2015). Therefore we emphasize the need for alignment between business and BA. As BA is a ramification of IT, we argue that the alignment between business and BA is a capability. COBIT (ISACA) identify the key enablers for BA alignment as being information, infrastructure, processes, policies and framework however this framework doesn't provide any practical recommendations on how to precisely reach alignment between BA and the objectives of the business.

4.3.2 The Alignment between Business Analytics and the corporate framework

To produce value from BA, the scope and the corporate strategy needs to be completely aligned with the organization's BA capabilities. Moreover, this strategy must be widely spread and implemented across the whole entity. All BA initiatives should be designed taking into consideration the needs of the end-users. Indeed, user-oriented projects consistently develop commitment and foster the usefulness of BA. This is why the organization's ability to foster the usefulness of BA to onboard its employees into following the pre-established corporate vision is vital for supporting the business needs of the entity. Besides, commitment from the corporate and operational management is also one of the main criteria that allow BA to be consistently integrated into the processes.

4.3.3 Business Analytics misalignment components

The development of the BA misalignment component is based on a system perspective. Therefore, just like ERP systems and IT, we consider that aligning BA with the needs of the organization is highly correlated with the company's capability to efficiently identify misalignment. To develop BA misalignment component we use the research developed by Van Groenendaal (2015) by reflecting the misalignment impacting ERP to BA (Table 5). This use of the ERP misalignment literature to create BA misalignment can be justified by the similar behaviour of these systems. Indeed just like ERPs, BA tools aim at capturing and processing data to provide an insight-based decision making. A misalignment creates a wide range of different consequences including the extraction of wrong data, the loss of trust in BA by stakeholders as well as the need for a change in BA tool usage. Also, misalignments impact all the layers of the organization which comprise, the operational, the tactical and the strategical levels.

Table 5. BA Misalignments based on ERP and Data quality literature.

BA Misalignments	Consequences	Layer
Poor data usability (dashboard)	• Incorrect data generation • Loss of trust in the business analytics tool the by end users	Operational
Poor stakeholders' visibility of BA calculation logic.	• Decision-making with questionable output • Lack of control about data output • Loss of trust in the BA tools by stakeholders	Operational
Contradictory input data	• Wrong input • Complex reports • Incorrect data • Wrong use of input data	Operational
Incompatible terms and meanings	• Loss of trust on the BA tools • Wrong input • Wrong use of input data	Operational
Poor system quality and performance	• Loss of trust in the BA tools	Operational
Incompatible IT infrastructure	• Loss of trust in the BA tools by stakeholders • Need for BA tools change	Tactical
Poor output quality or accuracy	• Wrong decision-making	Tactical
Complexity of reports and interface	• Wrong decision-making due to complex data	Tactical
Poor data reliability	• Misleading data reports • Wrong decision making	Tactical/strategical
Poor data relevance	• Impossibility to enforce decisions using data	Tactical/strategical
Complex data presentation and output format	• Incorrect reports • Inflexible reports	Tactical
BA tools conflicting with the organization vision and organizational structure	• Need for change in the business processes • Need for change in the organizational structure	Strategic
Incompatibility of BA tools with the business model	• Need for change of the BA tools	Strategic
Incompatibility of BA tools with and IT strategy	• Need for change of the BA tools • Need for change in the organizational structure	Strategic

4.3.4 Business Analytics misalignments resolution strategy

Based on the ERP literature we propose two solutions to correct misalignment between BA and the needs of the organization: (1) the customization of the BA tools according to the needs of the organization; (2) a change in the organization structure. On one hand, adapting the solution to the organization can be both costly and difficult to implement. On the other hand, it can be sometimes crucial for the organization to customize a certain feature of the system. Apart from those two opposed approaches, it can be more profitable for the organization both from a financial and structural perspective to use workarounds. A workaround consists in modifying the first intended use of the certain feature of the technology. Although coming at a low financial cost this solution can decrease the efficiency of the organization and limit scalability and integration. Moreover based on the IS literature we argue that to cope with misalignment the organization should create a mechanism that aims at identifying and resolving BA misalignment.

5 BUSINESS ANALYTICS GOVERNANCE

Within this chapter, three key concepts are defined: IT governance, data governance as well as BI&A governance. Because IT governance includes the concept of Data and Business Analytics, IT governance is defined in the first place. Also, because Data governance contains key elements for defining BA governance, Data governance comes after IT governance. Finally, we define the components of BA governance. In this chapter we emphasize that the concept of alignment is closely related with the concept of governance, and that governance represent a key enabler for reaching alignment.

5.1 IT governance framework

5.1.1 Definition of IT governance

As IT governance is a heavily studied topic, there are many definitions of governance. Weill (2002) refers to IT governance as the framework defining decision rights and accountability to foster desirable IT behaviour. On the other hand, Peterson (2004) suggests that IT governance also includes the process of monitoring strategic IT decision. Moreover, De Haes (2009) defines IT governance as a part of the corporate framework: "*IT governance consists of leadership and organizational structures and processes that ensure that the organization's IT sustains and extends the organization's strategy and objectives*".

5.1.2 The role of IT governance

Because the use of IT is developing and becoming a critical factor of success for the organization, the importance to govern and manage those processes becomes essential. Weill (2004) states that top-performing organizations are those that succeed to implement effective IT governance, therefore creating value from IT. Besides, De Haes (2005) emphasizes that the organization's capability to enforce IT governance is correlated with the organization ability to align IT with the need of the business. Nevertheless, the implementation of an IT governance does not obviously imply that the governance model is working within the organization. In fact, the design and implementation of an IT governance model finely tuned to the needs of the organization represent a challenging objective. Also, the implementation of a successful IT governance seems to be correlated with the use of the 3 following components (De Haes 2008):

- IT governance relational mechanism
- IT governance structure
- IT governance processes

The relational mechanism refers to the active collaboration and involvement of the relevant stakeholders. Moreover, the IT governance structure relates to the roles and responsibilities and the IT strategy. IT processes concern the IT decision-making process as well as the IT monitoring that can be supported by a variety of frameworks like ITIL or COBIT supporting the aforementioned process. Therefore, according to Weil (2004) the role of IT governance is (1) to assign the decision right to the correct groups to reach the desired objective and IT behaviour within the organization; (2) to clearly define the stakeholders accountable and responsible for decision-making; (3) to empower the managers in taking decisions without top-managers. Based on the importance of IT governance suggested in the above literature we conclude that IT governance is a key capability.

5.1.3 IT governance framework

Larsen (2006) realized a survey of literature that comprises seventeen IT framework aiming at establishing governance within the organization. Among these tools, COBIT and ITIL are the most used and accepted by organizations (Aguirar, 2018).

COBIT is widely accepted and emerged as an additional framework to the COSO (Treadway Commission's Committee of Sponsoring Organizations (Fedorowicz, 1998). "*With Sarbanes-Oxley (SOX) in the U.S. and other legislation enacted worldwide, effective governance over IT has become law for many companies"* (Hardy, 2006).

The COBIT framework created by ISACA is a proven set of standardized processes that businesses can be used to ensure that information technology is effectively and securely integrated with business goals" (Tusevski, 2011). The latest version of COBIT 5 is created using the former versions, although, this version emphasises more on processes. COBIT 5 framework is divided into two parts: management that evaluates planning, building, running and monitoring (PDRM) and governance that covers evaluation, direction setting and monitoring (EDM). Those two parts of the COBIT 5 framework aim at improving the performance of IT in supporting business processes (ISACA). Moreover, COBIT 5 is built upon seven enablers: (1) principles, policies and framework, (2) processes, (3) organizational structures, (4) culture, ethics and behaviour, (5) information,

(6) services, infrastructures and applications and (7) people, skills and competencies. Those enablers support the organization in optimizing its IT investment and resources (Bartens, De Haes, Lamoen, Schulte, & Voss, 2015).

5.1.4 Assessment of IT governance framework

Before choosing an IT governance framework, it is crucial to assess the current organization's settings which comprises the scale, the number of business units, the industry as well as the structure of the business units (Clementi, 2006). Moreover, the company's strategy should be widely understood across the organization.

As described above, there are different methods to deal with the organization IT governance. Although these methods can be lengthy and difficult to implement. To assess the current IT governance, a common method is to use the *"STOP: Strategy, Technology, Organization, People and Environment"*. Based on this approach the three following steps are developed: (1) define the core requirement of the IT governance according to the framework, (2) design the desired approach that combines the requirements, (3) illustrate the use of the approach (Bin-Abbas & Bakry, 2014). Moreover, this phase consists of identifying the key decision-makers as well as the contributors to IT governance. Also, the desired result of IT governance should be assessed using financial metrics and IT resources. Once the IT governance framework has been implemented, its successfulness must be monitored. To do so, COBIT provides an assessment model that relies on the maturity of the organizations' IT governance.

5.2 Data governance

How does data governance follows from IT governance

5.2.1 Data governance definition

According to Wende (2007), data governance consists in defining the roles, the accountabilities as well as the responsibilities to decision areas. Moreover, data governance is the foundation for the establishment of standards and guidelines that aim at ensuring that the organization maintains, collects, stores and presents qualitative data. Data governance and corporate governance must be closely aligned to ensure that the data leveraged by the organization complies with both its needs and the regulations.

5.2.2 The role of data governance

Former literature emphasizes the importance of using data governance to maintain the organization's data quality (Dai, 2019). Managing data embodies the foundation for the success of BA, therefore, this part outlines the importance of data governance and data management.

The role of data management is to deliver all the organization's stakeholders with the appropriate data. However, based on their position in the company, stakeholders have divergent data needs (e.g operational, tactical or strategical data). A corporate environment solely relying on data management for answering the data needs of each stakeholder can create difficulties and lower the quality of the data provided (Wende, 2007). Therefore, an integrated data management process is needed to answer the requirement of the organization on a business and IT point of view. While data management emphasizes on the *"collection, organization, storage, processing and presentation of high-quality data"*, data governance focuses on implementing an organization-wide *"accountability"* for data management (Huang, Lee & Wang 1999). In other words, data governance assigns responsibilities and roles to efficiently track the decision-making process while managing data. It also provides corporate guidelines and streamlines the data management practices to ensure the alignment between the organization's strategy and the high-level data rules in a wide complex environment (Wende, 2007).

Therefore, data governance has to be differentiated from data management (Bitterer & Newman, 2007) as data management provides a ”day-to-day” decision-making and data governance provides a framework for managerial decision-making. Taking into consideration that the research related to IT governance is more developed than the research on data governance, Wende (2007) chose to use the IT governance structure already existing to develop a model for data governance. This data governance framework is based on three main components: ”role”, ”major decision area”, and “assignment of accountability”. It is stated that IT governance and data governance should depend on governance principles. Moreover, the data governance framework should be closely aligned with business and IT data customers.

Because we showed above that IT governance was a capability and based on the correlation between IT and data governance, we conclude that data governance is indeed a capability.

5.2.3 Data governance critical factors of success

Cheong (2007) emphasizes that to ensure data quality the organization should focus on "*people, processes and technology*" and create a mechanism that assesses data quality. Moreover, Cheong (2007) shows that to sustain a successful IT governance framework, data governance is required. Marinos (2004) defines ten critical factors for successful data governance:

- The development of accountability for the leadership of the data governance process.
- The development of data standards making sure that the corporate data is aligned with its purpose.
- An alignment between data technology, processes and the business goal.
- The organization must acknowledge that data management is complex and ensure the coordination of data exchange between the relevant stakeholders.
- The data governance structure must cross the hierarchy and be structured to cover each layer of the entity. This permits to enforce the priorities and to foster the data quality.
- Metrics must be set to measure what data quality looks like.
- External third parties must be held accountable for the quality of the data they provide.
- Strategic data controls must be set to identify when and where data is controlled.
- Policies and procedures must be frequently controlled to make sure that they are correctly applied.
- Data users and stakeholders must understand the importance of data governance to develop their involvement in maintaining data quality.

5.2.4 Data governance enablers and inhibitors

According to Tallon (2013), data governance can be impacted by numerous organizational, industry and technological factors enablers and inhibitors. Concerning the organizational data governance enablers, it is suggested that the organization should focus on the business strategy alignment with IT and adopt a centralized IT structure. Although a decentralized IT structure, as well as a complex mix of products, has been shown as an inhibitor that complicates the adoption of standard and data policies. Furthermore, industry regulations can inhibit the implementation of data governance by obliging the organization to adopt an inappropriate IT structure. On the other hand, the data growth predictability can represent both an enabler and an inhibitor for data governance. In that case, if the industry factors do not provide data standards, it makes it complicated for the organization to foresee a potential structural change due to data growth. Finally, IT standardization and the organizational culture of promoting strategic use of IT are shown as being the key technical enablers for successful data governance. Nevertheless, IT legacy systems oblige the organization to handle multiple disparate systems and reduce the company's ability to protect and govern data.

5.3 Business Analytics governance

5.3.1 Definition of Business Analytics governance

As BA governance is a part of IT governance, we base our definition of BA governance on IT governance. BA governance defines the organization's capability to implement and maintain the business processes based on BA capabilities. BA governance ultimate goal is to provide the managers with the desired data to implement decision performance (Horakh, Baars, and Kemper, 2008). To achieve BA objectives, BA governance needs to be aligned both with IT and corporate governance (Grandhi, 2013). Therefore, BA governance tackles issues related to IT/business alignment but also, prioritization of BA projects, BA project management standardization, and data quality assessment and monitoring (Waston, 2007).

According to Fernandez (2008) BA governance refers to the procedure of designing and enforcing the infrastructure to align BA with the need of the organization.

5.3.2 The role of Business Analytics governance

BA tools are often very risky and expensive but can also bring a high return on investment. The issue is that many organization use BA tools on an ad hoc basis without clear planning on how to implement them (Heinrich, 2003). In some cases, the organization collects a considerable amount of data but misses the capability of processing this data into information useful to improve decision-making. Therefore, BA governance is critical for the company to implement and maintain BA on the long term basis.

As part of a BA project implementation, Fernandez (2008) emphasize that BA governance must be driven by four core values. The first value refers to the organization's capability to deal with changing the decision-making process. This value points out that to efficiently integrate BA within the organization, the entity needs to develop flexibility coping with changing information and business requirement. Thereafter, BA governance should fill the gap between business requirement and IT for instance by using multidisciplinary groups and system assessment by end-users. This will ensure the BA capabilities of the firm are aligned with the business needs. Also, BA governance must ensure the flexibility of the hierarchy, especially in project groups to foster information circulation along the life cycle of the project. This is why responsibilities and decision-making rights should not be fixed and change along the life cycle of the project. Finally, the project team in charge of the BA implementation must have a wide range of capabilities which includes project management, technical support and business support.

To summarize the role of BA governance is to ensure that the needs of the business will be correctly integrated into the system and that the set of tools will be finely tuned with the needs of the end-users. Finally, BA governance has to ensure that BA allows the business to maximize the benefits using resources in a relevant way (Muntean, 2013).

5.3.3 BA Governance framework

Analytics is currently being heavily developed but to fully harness the use of BA, the need for a BA governance framework emerges. Indeed, organizations are implementing advanced BA tools that need more topic related consideration that are not provided by Data and IT frameworks.

Avery (2015) emphasizes the need for a BA framework and proposes four key components that describe how to handle strategical data and address BA behavioural issues; (1) the accountability component provides the responsibility for analytics within the organization; (2) the accessibility component ensures that analytics and reporting functions are accurate, developed on time, available, and bring direct value to stakeholders; (3) the community component ensures that all business units fully participate with the business owners and the technical teams; (4) the uniformity component relates to the definition and implementation of policies for analytics reporting, key performance indicators as well as reporting cycle, information dissemination and periodic review for governance assessment.

Therefore, the analytics governance framework should be aligned into a system governance framework that is based on Data and IT governance. Such a framework would be divided into five sections:

- Data governance: where the data governance framework includes maintaining data quality.
- An IT governance: where an IT governance framework like COBIT 5 (ISACA) is implemented.
- BA governance: with a governance framework that is process-driven and defines the key performance indicators. This governance framework should answer the question: *"what the analytics is supposed to deliver"*.
- A decision making governance: where the concept of governance assigns accountability for decision-making.
- Result-sharing governance: where the governance fosters the communication across the organization of the result provided by BA.

Because BA governance is interdependent with both data and IT governance that have been showed as being capabilities. We conclude that BA governance is a capability.

6 A PRACTICAL FRAMEWORK FOR BUSINESS ANALYTICS

6.1 Research model

Throughout chapter 3, 4 and 5 we discussed the concepts of BA capability, alignment and governance. By creating this model we try to give a practical overview on all the dimensions that should be interconnected in order to provide the organization with an insight on how to reach performance using BA. Using the former chapter we justify the connection between each dimensions of the model.

This research model uses the model of Cao (2015) (Appendix 11) and displays the five critical dimensions necessary to reach BA performance discussed in section 3.3: BA technology, Data-driven infrastructure, Information process capability and Data-driven decision making and effectiveness. This model shows a positive relationship between decision-making effectiveness/Data driven-decision making and data-driven infrastructure. A data-driven infrastructure depends on the BA strategy and the BA structure of the organization as discussed in section 3.4. Avery (2015) states that to enforce a strategy and structure adapted to BA, the organization needs to implement a BA governance that is correctly aligned with both data and IT governance. The need for a governance dimension that monitors the data-driven infrastructure is therefore proven. This is why we add to the model the governance capability. The role of this dimensions is to enable a successful data-driven environment. BA governance should, directly and indirectly, support the five BA capabilities of the research model with the ultimate goal of developing and sustain the capability decision making effectiveness.

Because the components of the framework and the framework itself were assessed separately, the research model (Figure 2) only represents the dimensions of BA capability, the components of each dimension are thereafter included in the re-designed framework.

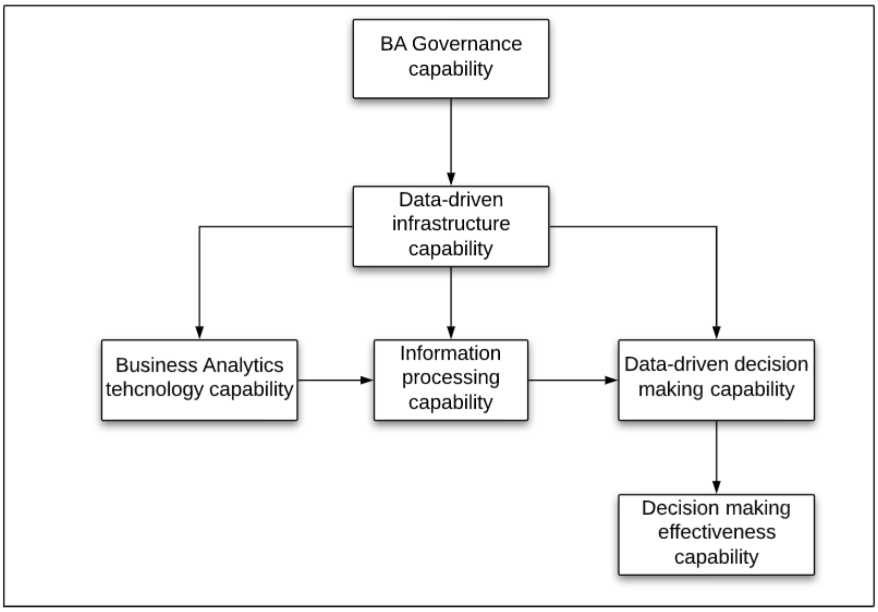


Figure 2. Research model

6.2 BA enterprise framework design

We note that all the components of the framework are described in section 3.3.

The re-design of the research model results in a BA enterprise framework that maps the capabilities and components necessary to achieve the final capability: decision-making effectiveness (Figure 3).

Grembergen (2005) demonstrates that BA governance has a positive effect on the BA/business alignment (chapter 5). However, El-Tebany (2014) shows that to reach alignment and to cope with misalignment, a governance capability must be incorporated into the organization's strategy (section 4.1.3). As described in the fourth chapter, BA alignment is also a capability but results from the creation of a BA governance (chapter 5). Therefore BA alignment is here considered as a sub-capability of BA governance and embodies the organization ability to enforce alignment as well as its ability to create a mechanism to cope with misalignments as discussed in chapter 4.

Therefore, we incorporate to the model of Cao (2015) the BA governance dimension that includes the component (1) BA alignment with the organization strategy but also a (2) BA mechanism that aims at identifying and resolving BA misalignment.

Analyzing chapters 4 and 5, we can state that the role of BA governance is to align BA with the business objectives and to identify BA/business misalignments. Based on this BA governance capability definition, the organization can integrate BA within its strategy, its structure, its policies and rules and its business processes which creates a data-driven infrastructure capability. As stated by in section 3.4 a data-driven environment represent a foundation for correctly managing both information processing and BA technologies. Furthermore, the BA technology directly influence the organization ability of creating an information processing capability and directly impact the data-driven decision-making. The information processing dimension defines the organization ability to provide decision-makers with data following the needs of the organization in terms of data quality as described in section 3.1. The data-driven decision-making dimension is influenced by the three previously mentioned streams(data-driven environment, data processing capability and BA technologies capability), but also indirectly relies on the BA governance capability.

6.3 Pre-requisite for using the BA enterprise frameowrk

The organization's ability to reach a BA technology capability is not related to the use of a defined set of tools but to its ability to fulfil its business needs leveraging the right technologies. To use this framework, the organization must first define the BA tools necessary to reach decision-making effectiveness (Cao, 2015). The technological tools part of the framework are uniquely related to the organization investigated during the case study. To use this framework to assess a different organization, the BA tools of the given entity must have been previously assessed and selected. Therefore, the BA technology needs will differ according to the goals of the organization.

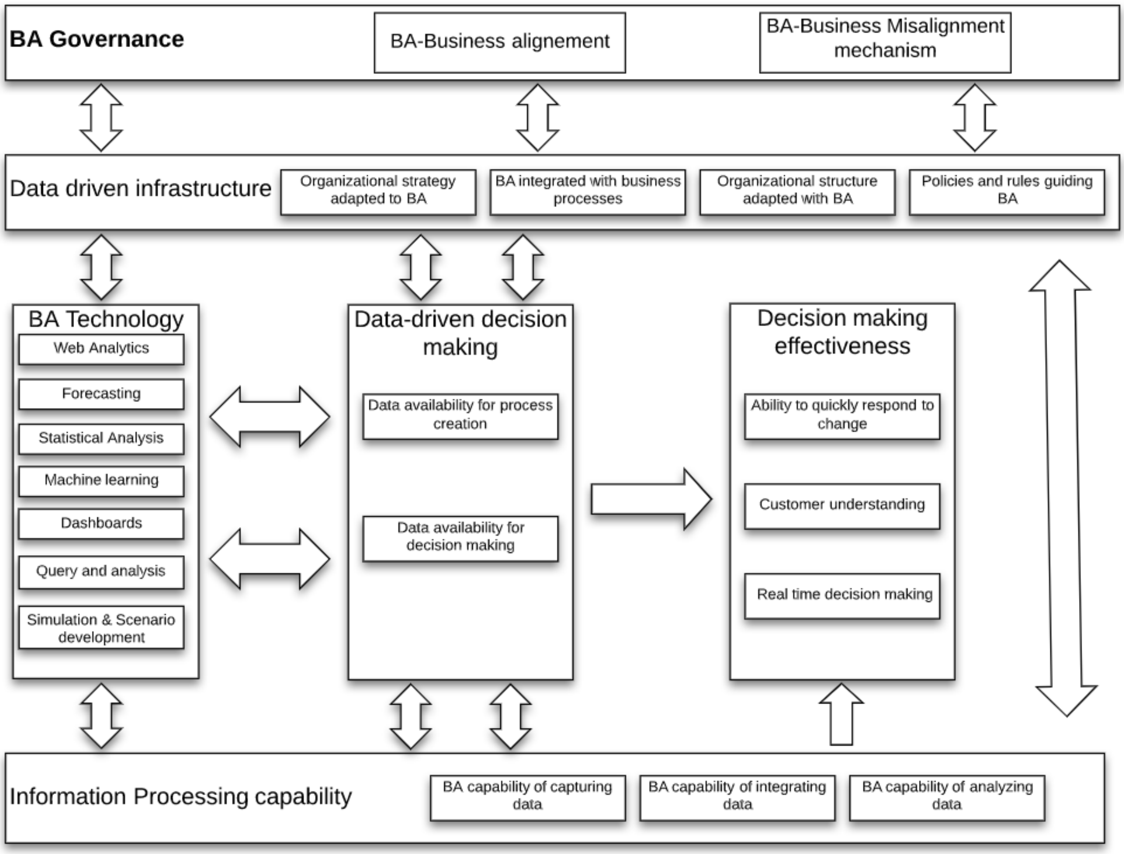


Figure 3. Decision-making framework

6.4 Framework validation

The experts from Amazon were asked to give their feedback on the different components of the decision-making framework. The feedback provided by the experts did not change the model itself but provided an additional definition to each component, therefore, completing the scope of each criterion.

6.4.1 The organization structure is adapted to Business Analytics

The organization structure defines how Business Analytics activities are divided across the organization and highly impact the efficiency of BA. Each business units should have BA available skillset to provide managers and directors with the data needed for making decisions. The structure of the organization should be designed to enforce a dynamic exchange of metrics and data across the organization. Although, the organization should provide precise rules concerning how data should be exchanged to protect the customer. A good data structure also fosters the rapid communication of data from the operation layer to the strategical layer.

6.4.2 BA is fully integrated within all the core processes of the company.

In order to identify process inefficiency and breakdown, the organization should be able to retrieve data on time at each stage of the process. Also, this data should be efficiently communicated throughout the business units to rapidly find the root cause of the problem. Besides, the process owner should be able to rapidly and flexibly challenge sub-efficient processes using BA.

6.4.3 The company policies and rule guide BA

The policies and rules should enable data to circulate at a fast pace within the organization but also ensure that the data is correctly maintained and updated according to the needs of the business. Also, it is important to create a Service Level Agreement with data owners to make sure that the data needed to make a decision is always available when needed. Data security and data privacy should also be heavily fostered by the policies and rules of the organization.

6.4.4 The entity is capable to efficiently capture data.

This capability relates to the organization ability to capture internal data related to the business processes but also external data related to the partners and customers. This capability is the first step of the process but when dealing with a huge amount of data, organizational data ownership and access can become very tedious. It is very important to make sure that data that has been captured is unaltered and stored in the right table in a timely fashion. To develop this capability, the organization should be able to efficiently forecast the data growth.

6.4.5 The entity is capable to efficiently integrate data.

Data integrating refers to combining data from different sources (business processes, external data) in the right location. Once the data is correctly integrated and stored, it can be used by the data client. Integrating a huge amount of data can be very complex especially when the data sources and types are very diverse. This capability is key because the organizations capture a huge amount from different data sources,

6.4.6 The entity is capable to efficiently analyze data

This capability relates to the organization ability to model semi-refined data to provide useful information. The data analysis complexity highly depends on the data captured and the integration steps. The location, the size and the sanity of the data impact the organization ability to analyze the data. Managers should ensure that technical employees have access to well-structured, qualitative data.

6.4.7 The entity is capable of almost exclusively rely on data for decision making

This capability highly depends on the culture of the organization. The organization structure should be designed in such a way that employees can make every decision using qualitative, task-related data. The organization’s capability to exclusively leverage data to make decisions is crucial not only for directors and managers but for all employees. This implies that insight-based decision-making should be fostered and enabled at each level of the hierarchical structure.

6.4.8 The entity is capable of using Business Analytics for innovating

The capability relates to the organization use and strategy for BA. To innovate, the organization must step out of its day-to-day use to identify new processes. To innovate, the management side must be inclined to allocate resources and foster interaction between top-managers and the technical side must be dynamic. Also, managers should have a relatively good understanding of the technical challenges.

6.4.9 The entity uses Business Analytics to challenge current practices

This capability relates to the organization's ability to create strong bottom-up communication between employees leveraging BA and the managerial side. To challenge the current practices managers should first be aware that a business process is sub-efficient. A critical factor for the development of this capability is the organization's ability to foster an intensive communication intra- and extra-business units.

6.4.10 Business Analytics technology

The technological capability relates to both the presence of the necessary technology within the organization but also human skills required to implement them. Creating a BA technology relies on hiring and training the best talents but also to transfer the knowledge within and outside the business unit. From a tool perspective, the management side must consistently understand the needs of the business before implementing a tool. Moreover, the tools roll out must be planned and its expected benefits must be previously defined.

7 CASE STUDY AT AMAZON

7.1 Company description

Amazon, established in 1994, is an American multinational organization based in Seattle and focuses on cloud computing, e-commerce, artificial intelligence. The organization is recognized for being disruptive using technological innovation. It is considered as one of the big four technology organization together with Apple, Microsoft and Google. It is currently the world most valuable company with annual revenue of 280.4 billion dollars in 2019 and counts more than 840 000 employees in 2020. Amazon owns more than forty subsidiaries in various sectors including innovative technologies, food and aerospace engineering. In 2020, the company sells approximately 4000 items per minutes worldwide. Amazon started as an on-line seller distributing mostly books before diversifying its offer to a huge variety of products including product of almost any kinds. Finally, one of the most compelling figures is the Amazon revenue per user that is approximately 189 dollars (6 times greater than E-bay).

One of the organization greatest innovation was its personalized recommendation system that has been created on top of the big data capability of the company. Concerning its core business model, Amazon has always based its strategy on a very thin margin always prioritizing customer experience and low prices over high-profit margins. In 2020, Amazon’s strategy has not changed as emphasized the COVID-19 situation. Indeed, instead of taking advantage of the pandemic situation to increase its profit, the organization has chosen to once again bet on customer experience to increase its 43% of market share. Indeed, instead of only collecting the benefits brought by the huge increase in online product demand during the pandemic, the company has chosen to upgrade its services by hiring 100 000 employees only in the US.

Concerning its structure, the organization is characterized by its bottom-up hierarchy. Nevertheless, unlike much other hierarchical organization structure, the company developed high flexibility to cope with its changing environment in each market place. Taking into consideration that Amazon has more than 800 000 employees and that each market places tactical strategy is different from one another, applying the model to the entire company would not be reliable as it would require a much wider range of experts.

Therefore, this study will only focus on one department. The department part of this case study focuses on maintaining the operational part of the business by monitoring the processes and preventing process break down. Also, an important role of the department is to implement new projects that will improve Amazon's customer experience as well as its efficiency.

7.2 Business Analytics at Amazon

Concerning its data capability, Amazon is well-known for being one of the most data-driven organizations in the World and developed a strong culture of evidence-based decision-making. The data capability of the organization represents one of its greatest competitive advantage and products like Amazon Redshift (cloud data warehouse) are both used internally and proposed as a service for external third party. To cover its heavy investment in data storage, analytics, security and application integration, the organization has chosen to propose its highly sophisticated data leveraging tools and systems on the market. In terms of algorithm, Amazon developed A9 which is the system that is used to rank the product in a search result. This system is similar to Google search engine as it processes keywords to decide what the most relevant search result are. Nevertheless, unlike google's, the A9 algorithm strongly emphasizes the sale conversions promoting the products that are most likely to end up in a sale. Therefore, the A9 algorithm will prioritize products that have are more likely to receive high traffic.

7.3 Business Analytics Misalignment investigation

In this section, we define the BA misalignment investigated throughout the research and internship at Amazon. Moreover, in section 8.2, we link the identified misalignment to academic research to provide recommendations.

7.3.1 Misalignment identification

The soft data collected during the interviews as well as project observation within the department leads to the identification of a set of misalignments within the company. Indeed, for each component, experts were asked to justify their grades. Based on the experts justifications given on the current state (as-is) of the department, we have gathered the misalignments that came out more than once during the interviews. Moreover, we investigated those misalignments throughout project observation, documentation and implementation. Based on the BA misalignment components defined in section 4.4.3, captured misalignment within the department are listed (Table 6).

Table 6. Amazon's Business Analytics misalignment

Component	Misaligment	Effet on the business
Business Analytics	a. BA is mostly used to monitor the business.	The BA impact on process creation and innovation is limited.
	b. Table for query and analysis are missing.	Missing attributes make SQL query writing very tedious which create errors and is time-consuming.
	c. Dashboards tools don't support the hight amount of critical metrics necessary for external stakeholders to understand the department challenges.	Decrease the efficiency of bottom-up communication with directors.
	d. Machine learning potential is not used at its full capacity.	Algorithms are sub-efficient which considerably decrease operational productivity.
Information processing capabilities	e. Data ownership is unclear.	Sanity checks are often required. Leads to mistakes in pattern creation and limits data-driven decision making.
	f. Data schema are not always maintained.	This creates problems in term of data quality which creates data pattern errors and inefficient time-consuming workaround.
Decision making effectiveness	g. BA and change management are not aligned.	This creates latency between BA and the business change which creates process inefficiency.
	h. Poor data velocity.	Limit real-time decision making
	i. Various sources of diverging data.	Limit the impact of data-driven decision making.

7.3.2 Misalignment contextualization

(a) First, we highlight that all Business Analysts and Data Scientists stated during their interview that monitoring the business and tackling day-to-day process break-down was taking the huge majority of their working time. It was also emphasized during project observation and interviews that, while monitoring the business, experts did identify potential process improvement.

However, even if those potential business process improvements are communicated and acknowledged by managers, they are put in a queue due to time constraint and are not implemented in a timely fashion.

(b) The absence of table critical for query-handling efficiency are missing. Indeed, as part of project implementation, observation but also in during interviews with Business Analysts and Data Scientists, we identified a key column missing from all critical tables. To tackle this issue, a huge query section based on a SQL CASE statement of more than 500 lines of code was used to include the missing attributes in the weekly reporting but also in a big majority of the queries. This workaround consistently impacted the time spent on running and writing queries but was also prone to errors. This misalignment creates a more high-level misalignment: (h) poor data velocity.

(c) The considerable amount of metrics needed to monitor the business made it difficult for the department to organize the dashboard in a meaningful way. This misalignment was identified while using the dashboard and was a problem stated and documented from the beginning of the internship at Amazon. The lack of metric overview using dashboards impacts the communication with top-managers and therefore limit metrics sharing which, in turn, consistently reduces communication efficiency and decision making.

(d) During project implementation, it was observed that a critical query for the business was sub-efficient and produced costly inefficiencies. Nevertheless, it was described by a Senior Manager from another department that machine learning was an obvious choice for improving this query. Although the department had the skills and technology to improve the process using machine learning the query had not been improved during the last two years due to time constraint. We note this project was implemented during the time spent in the department and was successful.

(e) Due to the huge amount of schemas, tables and databases, the data ownership within and outside the department is somehow unclear. Even if each database and booking table is owned by a specific department, the tables are very often used by a huge amount of users which makes it difficult both to assign the right (write and read, write and edit), but also to update and maintain those tables. Also, table updates are very often not communicated to the relevant stakeholders which creates huge bottlenecks and process breakdowns. Also, the lack of maintenance on some tables provides the users with unreliable data (f).

7.4 Expert interview

This section compare the results gathers throughout the expert interviews. This section support the section is the foundation for a gap analysis in section 7.5

7.4.1 Rating component importance

During their interviews, the experts were asked to rate the component of the model based on the importance that each of them has for the company (Figure 4). The data available to understand the customer is the component with the highest importance which is closely related to the *"customer obsession"* vision of Amazon. In our case, the BA tools technology is ranked the lowest even if the query analysis tool has high importance for the department.

Data processing overall dimension is the section with the highest importance which emphasizes the fact that technology is much less important than the organization capability to harness tools to create consistent decision-making.

Figure 4. Experts component rating importance

7.4.2 Rating component importance current state

Thereafter, the experts were asked to rate the same components but based on the current situation of the company (Figure 5). The experts rated the data availability for deciding within the department as the most important dimension. BA technology dimension obtained the lowest rank and the data-processing dimension the highest. Herein, web analytics and machine learning are two BA tools particularly low-ranked which can be explained because those components are also represented as relatively important.

Figure 5. Expert current situation rating components

7.5 Gap analysis of Business Analytics at Amazon

This section's objective is to identify the gap between the to-be-achieved state of the organization and its current situation. In the first place, we compared the results for the four major dimensions: data-driven infrastructure, data-processing capabilities, data-driven decision-making/decision-making effectiveness BA technology. After this, we compared the components part of each section.

7.5.1 Dimensions overall gap analysis

The first finding is that the BA capabilities current state within the company is very close to the expert expected to-be state. This shows that the BA capabilities of the organization have almost reached stakeholders expectations. Nevertheless, the BA tools and decision-making dimensions still have room for improvement, this reflects the problem communicated by the experts during the framework creation. This gap between the BA capability dimensions and the two other dimensions comes from various factors related to maintaining the data. Indeed, the size of the organization makes it difficult to track the ownership and sanity of both refined and raw data, which sometimes create difficulties using BA tools and developing efficient decision-making on a real-time basis.

7.5.2 Business Analytics gap analysis dimension

The major gap within this dimension concerns web analytics, this can be explained by the fact that the department focuses much more on the operational level (Figure 6). Therefore, the use of web analytics is relatively relevant as the function of the department is mostly to improve business processes. The output needed from web analysis is brought by other business units of the organization. The gap between the expected state of the BA tools capability and the current state is relatively low but still reflects a need for improvement in components like machine-learning and foresting. Those two gaps arise due to the lack of experience the department has managing those tools.

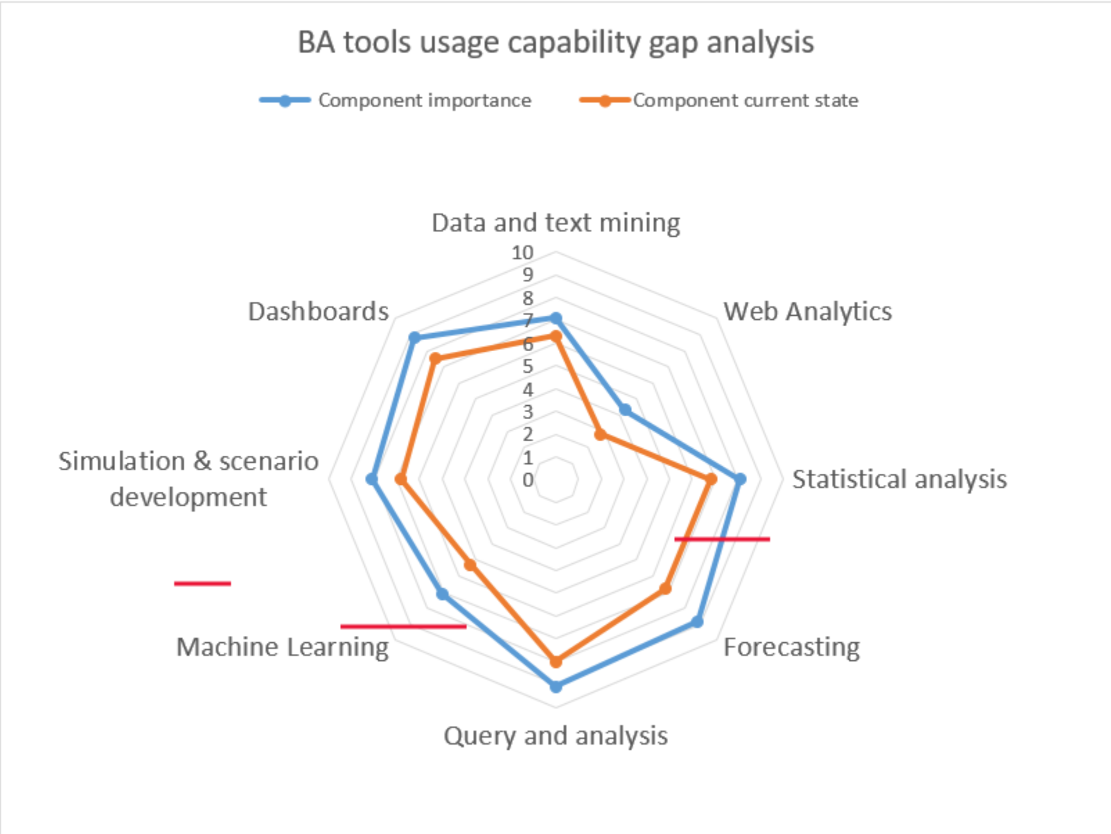


Figure 6. BA tools gap analysis

7.5.3 Data-driven infrastructure capability gap analysis

Analyzing the BA capability of the department, we could notice that, here again, the desired situation is close to the current situation. Although, the strategy of the department guiding BA component shows the potential for improvement. Indeed, according to the experts, the BA strategy is mostly focusing on monitoring and maintaining current business processes but less in creating new processing leveraging BA. Concerning the policies and rules guiding BA, it has been pointed out that the gap between desired situation and the current situation was mostly related to the lack of data ownership and data maintenance.

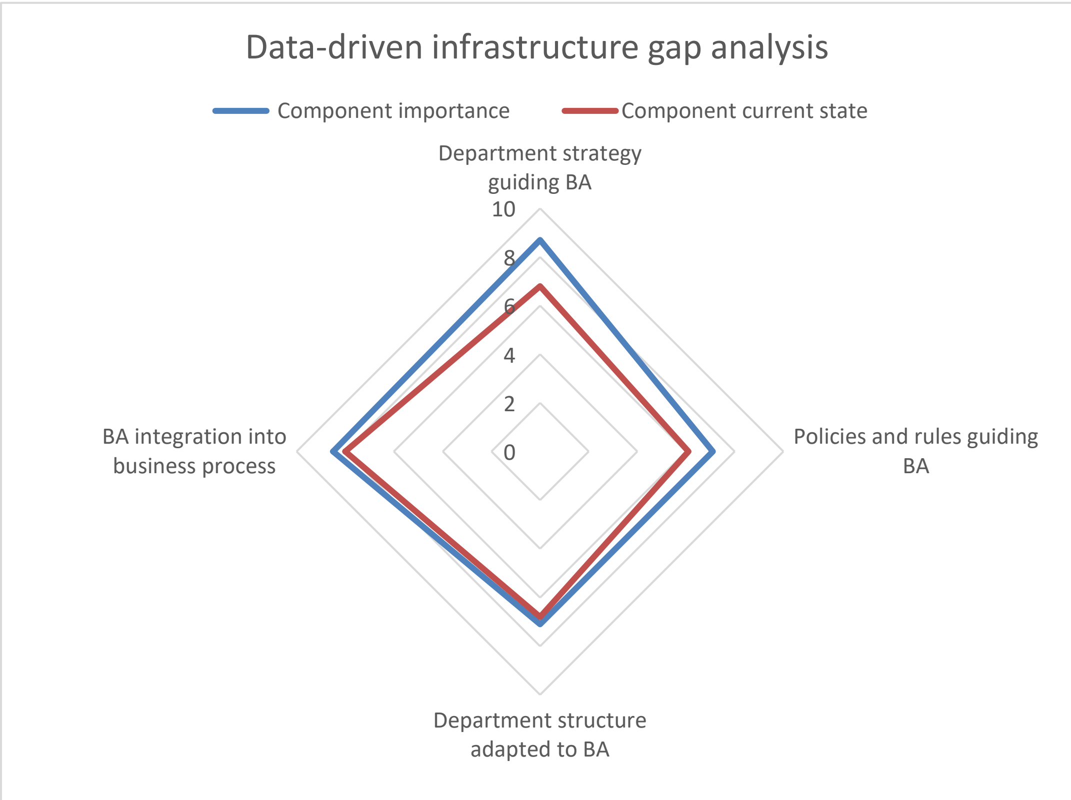


Figure 7. Data-driven infrastructure capability gap analysis

7.5.4 BA data-driven infrastructure capability gap analysis

The BA capability of integrating data has been described as sub-efficient due to the lack of ownership and schemas maintenance. Although, it has been noted that the department capability to collecting data was extremely efficient even if the important amount of tables made it difficult to rapidly locate and analyze data. These statements are supported by Figure 8.

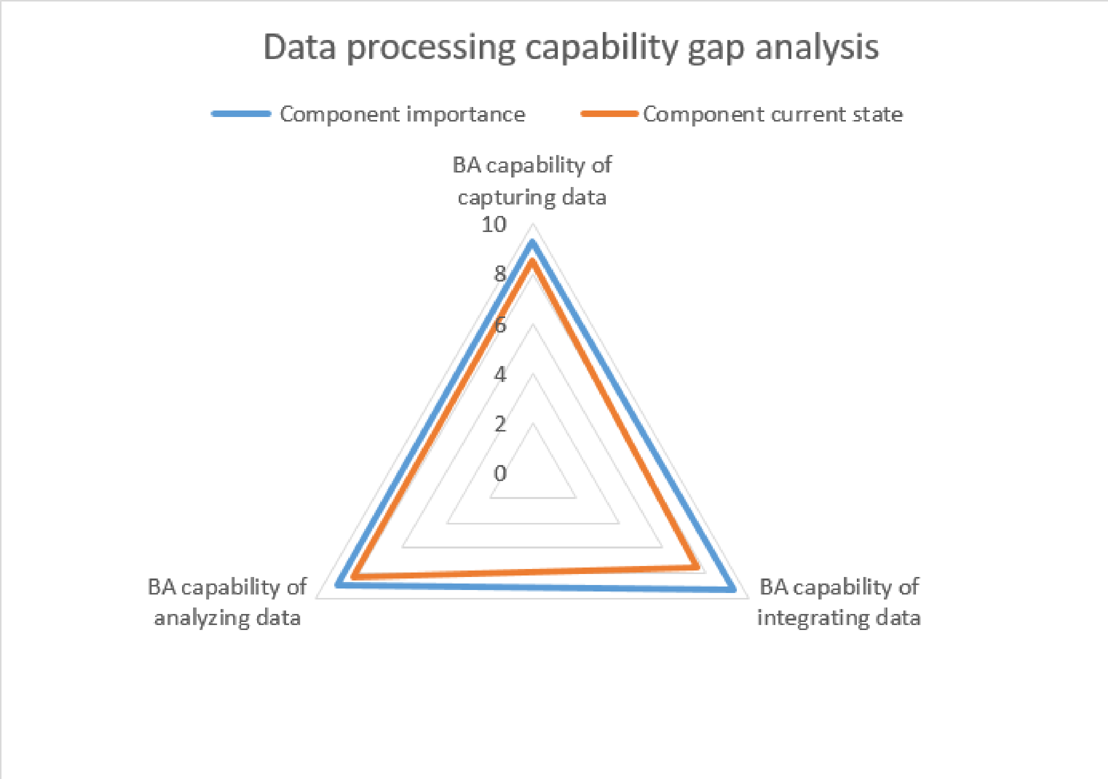


Figure 8. Data processing capability gap analysis

7.5.5 Decision making capability gap analysis

As shows the Figure 9, the two biggest gaps regarding the decision-making capability of the department are real-time decision and quick response to change. As described in the problem pointed by the experts, the data necessary to implement a change is present in the company but difficult of access and sometimes requires a sanity check. Also, the high number of critical metrics within the department makes it difficult to efficiently use data visualization tools to make a decision in real-time. Concerning the other components, it seems that BA's role of supporting decision-making is mostly achieved as the desired to-be situation is close to the as-is situation. For instance, experts clearly stated that data required to monitor the business was *"almost always present within the organization"* but sometimes tedious to find and process.

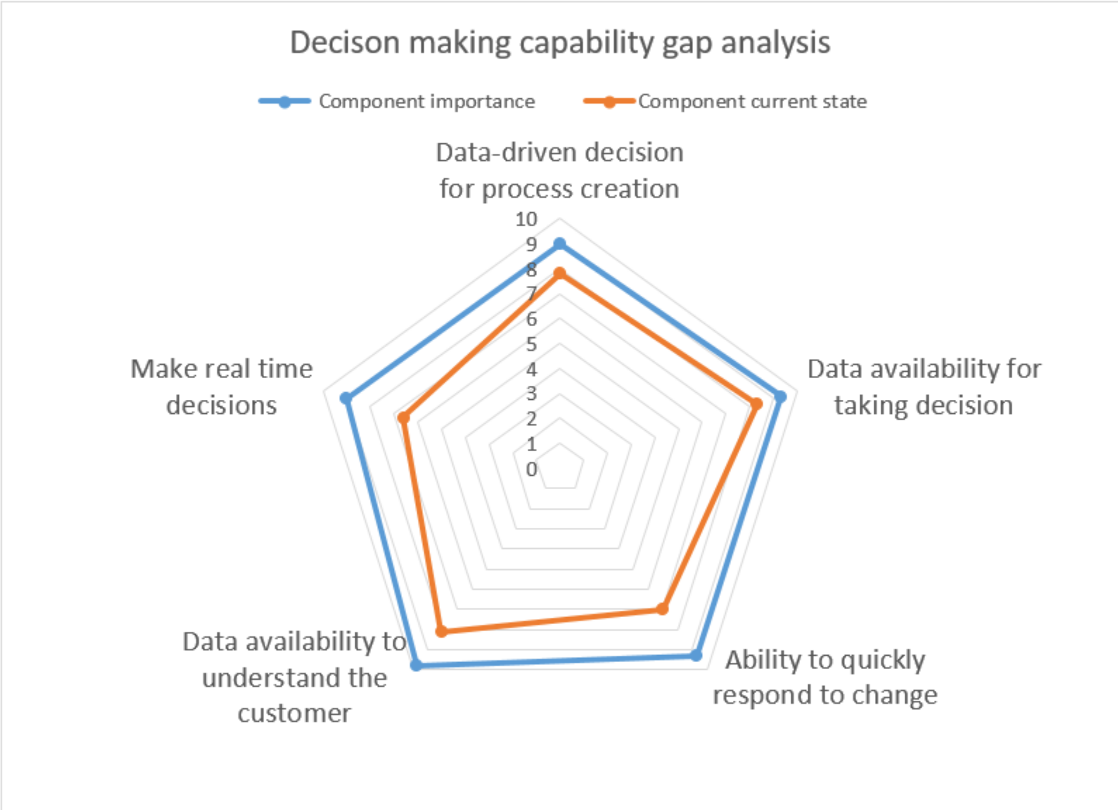


Figure 9. Decision-making gap analysis

8 RECOMMENDATIONS FOR AMAZON

In the first place, we selected the components that have the most important gap between the ideal situation and the current situation. Then, we linked those components with the academic literature to provide recommendations. As formerly mentioned, the recommendation is only made for the components part of the dimensions data-driven decision-making/decision-making effectiveness, data-driven infrastructure and data-processing and BA technology capabilities. In the second time, we used the same process to provide recommendations for the misalignments listed in section 7.3.

8.1 Gap analysis recommendation

In this section we provide an overall gap analysis in order to point out the component that most need to be improved (Figure 10). The below figure captures the gap between current and the desired situation at Amazon. Based on this figure we provide the recommendation.

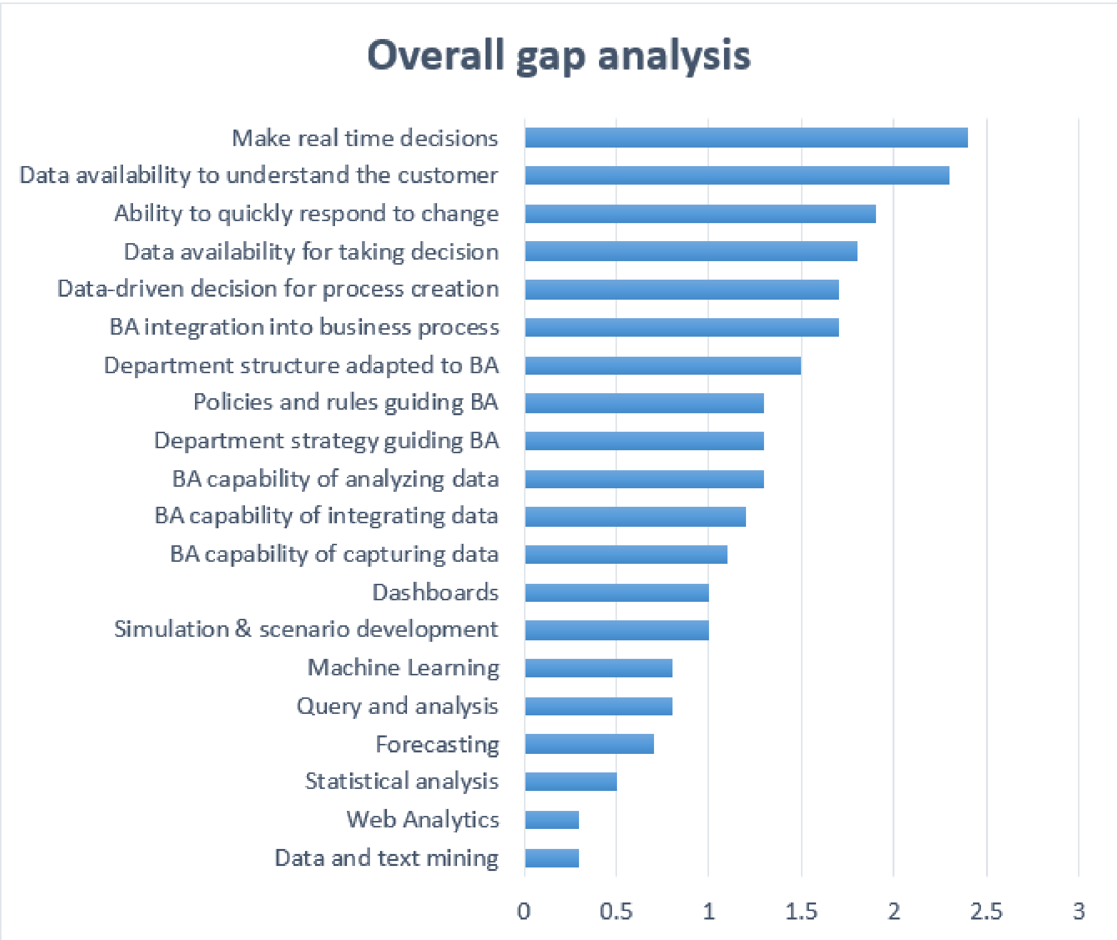


Figure 10. Overall gap analysis

8.1.1 Real-time decision-making capability

To successfully implement and monitor BA to provide real-time business intelligence, the entity has to overcome structural and technical challenges. In the first place, it is important for the organization to define what is a real-time decision and to adapt this definition to the needs of the business.

Moreover, to reduce latency impeding real-time decision, the organization must gather three factors: technical solution, a process change and a strong business case to involve the key stakeholders.

Also, an important component stands in constructing a flexible architecture that will scale as the need for data velocity grows. Furthermore, (Watson, 2007) shows that the entity must first help BA users to understand the potential of real-time decision-making. Once users understand the stakes and the possibility of implementing real-time BA, it increases their involvement in sustaining and implementing the capability. From a technical perspective, real-time BA necessitates the automation of the Extract Transform Load (ETL) processes to minimize as much as possible the human input except when the system monitoring the ETL jobs detects an error. Besides, the key queries at all level of the organization must be monitored to co-exist.

Finally, entities that tend to implement real-time decisions need to adapt their management practices and involve decision-makers to grasp the benefits of the capability.

Therefore, creating a clear definition of real-time decision-making adapted to the department, a flexible structure and users' and stakeholders' involvement sets the foundation for achieving real-time decision making.

8.1.2 Strategy guiding BA

The component strategy guiding BA objectives is to create a data-driven environment and is therefore critical for the successful implementation of BA within the entity. The strategy guiding BA reflects the entity's capability to develop a clear strategy and policies to enable the integration of BA within the organization's business processes.

According to Cao (2015), in order to create data-driven decision-making, the entity must produce new policies that are closely linked with BA. Also, these policies must be clear, easy for users to understand and facilitate the use of BA. Also, it is crucial to create a relevant organizational structure so that BA can be integrated into the organizational practices and therefore improve the decision making process and outcome.

Also, the allocation of BA human and technical resources without a clear strategy may be blurry and lead to overtaking the budget or to poor BA capabilities.

Finally, the absence of the component might lead to unclear objectives where the entity loses track of the goal of the BA capabilities.

These factors consistently impact the entity's ability to create a data-driven environment. Therefore, for the department to use BA to the fullest, it must improve its data-driven environment by implementing specific BA policies and strategy as part of a BA governance mechanism.

8.1.3 Data availability for customer understanding

This component is closely related to the concept of customer knowledge management which refers to *"knowledge creation, dissemination, acquisition, representation, sharing and utilization"* (Chan, 2004). Customer knowledge management addresses the customer knowledge collected throughout different sources: Customer Relationship Management operation, big data or external parties. Because the department objective is mostly to maintain and improve operational processes, most of the metrics used relate to internal processes. Therefore, the department as a reduced visibility on the impact of the business processes on customer satisfaction. By linking its metrics directly with customer satisfation, the department could produce a better customer understanding.

8.1.4 BA Capability of integrating data

An important amount of tasks performed by BA users consist of interacting with various information sources scattered across and outside the organization. To retrieve the desired information the user must know the databases structures and contents. Often generating the desired information requires breaking down a given task into a sequence of queries handling temporary storing. Although breaking down the task into a sequence of queries is required for ad hoc analysis, for task related to daily or weekly metrics needed to monitor the business it may be more efficient to create a secondary table that gathers intermediate data (Arnes, 1993). This intermediate data table reduces the daily query processing time and provides more flexibility accessing the information.

8.2 Business Analytics misalignments recommendation

8.2.1 The change management is not aligned with BA

In this case, the lack of change management related to the misalignment refers to none-strategical changes. Therefore this misalignment embodies a surface misalignment not related to the structure of the organization. This is why we only provide an operational recommendation. In this case, the change management is not aligned with BA, in more precise term, the tables are not always updated following the change of a business process which create incorrect data collection.

Kanter (1992) states that change management can be enhanced by improving intra and extra business unit communication. The successfulness of change is also based on involving all stakeholders into the decision making process and on the clear definition of the leadership. Also, as stated by Kanter we recommend the implementation of specific policies that link process change with the table to updated.

8.2.2 Data schema are not always maintained

In order to avoid poor data quality due to the lack of maintenance, Wende (2007) suggests to implement both a data management structure and a data governance. The data management structure ensure that the data is organized and maintained.

Moreover, data governance should ensure that the role and responsibility for maintaining the data are well defined. Therefore in this case a more clear data management must be implemented in order to ensure that each table is maintained over time.

Finally, it also seems relevant to review the data governance as the lack of schema maintenance is also due to unclear roles and responsibilities concerning the ownership of the scheme.

8.2.3 Dashboard misalignment

The most frequent issues that organization have to deal with while designing a dashboard includes the tool misalignment with the organization objectives and the entity lack of clear goal concerning the usefulness of the dashboard. Therefore, the lack of capability many companies have to correctly implementing this critical tools has an important impact on the organization's ability to efficiently monitor both the operational and strategical business objective. The core factors to an effective dashboard implementation project includes a clear understanding of the audience and the preparation of the key business metrics. Moreover, the ease of use and interface flexibility of the tool as well as the "sustained leadership" using a dashboard champion have also been shown as criteria that improved both the audience and sender commitment in using the tool (Kawamoto, 2007).

The definition of proper metrics can be enabled by producing an initial focus on the organization objectives. The defined metrics have to be meaningful for the business, assess the results and provide a precise insight on the further actions to be taken. Therefore, while starting a dashboard, focusing on a few key metrics that will ensure success and strategical insight seems to be the best approach. From an executive support perspective, initiating a dashboard requires the support of a top manager that understand completely the business and has sufficient authority to enforce the change. Without this top management support, the dashboard may fade away due to a lack of relevance. Finally, the last criteria relates to the simplicity of the dashboard, the tools should be easy to implement and to use, also, it should provide with an integrated range of information (operational, financial and initiative specific) (Kawamoto, 2007). Even if starting with few metrics is recommended, a strategical dashboard that displays complex metrics like IT business contribution represent a goal to be achieved in the long run.

9 CONCLUSION & LIMITATIONS

Along the past two decades, BA has become a competitive advantage for both process and customer understanding. Although, former literature has shown that BA tools are too often implemented without a precise vision and governance structure. This considerably impacts both the benefits brought by the technology and the involvement of the stakeholders in consistently trusting BA. The purpose of this thesis is to provide an understanding of the mechanism throughout which the organization can better yield the benefits of BA.

In the first place, we defined what is BA and what are the key concepts that it embodies and we defined most of the tools that are investigated throughout the case study. This literature review answer the first question. Moreover, to answer the second question, a definition of IT capability is produced and used as a foundation for proposing a definition of BA capability. Thereafter, we defined the core dimensions of BA capabilities. To answer the sub-research questions, we explain how these capabilities are built and what are their main components. Even though each dimension has its own components related to its success, we introduced the concept of BA governance that embodies the foundation for linking the different dimensions into achieving successful decision-making.

In order to answer the third question related the concepts of alignment and misalignment, we use the concept of IT alignment and misalignment to define what is a BA misalignment. Also, we argued that to reach alignment, an organizational mechanism aiming at identifying, resolving and preventing those misalignments is needed. Using ERP literature, a set of BA misalignment tools is created and solutions for resolving these misalignments were proposed. Using the IT literature, we answered the second part of the section by validating that the organization's ability to align its BA systems with the business is a capability. We also used the same logic and validated the fact that BA governance is a capability. Herein, we emphasized that the alignment capability is a sub-capability of governance. Concerning question five, we explained the theoretical validity of the research in section 1.4.4 and use academic literature to support the framework design validity. Finally, by providing expert feed-back and approval of the framework we develop the empirical validity of the research framework in section 6.3.

The creation of a case study permitted to apply the previously designed framework to Amazon. As part of this case study, we assessed each component of the framework within the organization using the experts' inputs. Thereafter, we created a gap analysis based on the differences between the current state of a component and its desired state. Then, we assess the components having the highest gap and provide recommendations to Amazon accordingly. Based on the description of the method used to apply the framework to Amazon, question six is answered.

To conclude, by providing the organization with an insight into the key BA improvements needed to reach a fully developed decision making effectiveness, we provided the organization with the foundation for the creation of a roadmap. Also, the model can be used for increasing the knowledge of the organization stakeholders. Therefore by creating and applying the model we answer the main research question.

9.1 Limitations

One of the most critical limitations of this study is related to the validity of the framework. Indeed, the fact that the framework is only validated based on experts feedbacks decreases the validity of the framework.

Also, due to the time constraint, the case study has only been applied to one company which considerably decreases its external validity. Moreover, the framework has been only tested within the department which increases the internal validity but decreases the external validity as all the expert worked on the same operational environment.

Furthermore, a limited number of experts have tested the framework which limits its wider validity.

Finally, the recommendations are only based on academic literature which decrease their usefulness from a business perspective.

9.2 Results versus literature

Because there is very little comparable research, confronting the results with the literature is difficult. Although there is research that defines BA capabilities, no research link BA capability with the concept of alignment and governance. However, the use of ERP misalignment as a baseline for a field research has shown to be rather successful as several concepts used in the ERP misalignment literature seem to apply to the BA misalignment.

For instance, the workaround has been shown to be a commonly used technique used by employees to cope with a given misalignment.

9.3 Wider implication

This framework can be used on other organizations but to use it, the organization must first assess its needs in term of BA technologies as stated in chapter six. Moreover, this framework is a very high-level framework which provides an overview of the BA capabilities. Therefore, that framework may not be applicable for resolving more precise and practical BA issues. Thus, this framework is more suitable for corporate organizations or for organizations having extensive use of BA.

9.4 Further research

To enhance the validity of the framework, we recommend developing a more extended set of tests with different organizations having a diverse use of BA.

Also, conducting interviews with a larger number of experts to validate the framework could enhance the business scope of the framework.

Finally, we recommend to extend the empirical validity of the framework by applying the framework, for instance, during the implementation of BA governance framework.

REFERENCES

Chan, 2004. Big Data Customer Knowledge Management. Communication of the IMMA: Vol. 14: Iss. 3, Article 5.

G. Premkumar, K. Ramamurthy, and C. S. Saunders, "Information Processing View of Organizations: An Exploratory Examination of Fit in the Context of Interorganizational Relationships," Journal of Management Information Systems, vol. 22, pp. 257-294, 2005.

Fernández-González, J. (2008). Business Intelligence Governance, Closing the IT/Business Gap. UPGRADE, 1(9),23-30.

Indranil Bose, Radha K. Mahapatra (2002). Business data mining – a machine Learning perspective. Information & Management 39, 211-225.

Kurniawati, Shanks, Bekmamedova (2003). The Business Impact Of Social Media Analytics.Association for Information Systems AIS Electronic Library (AISeL)

Groendaal, Beijsterveld (2015). Solving misfits in ERP implementation by SMEs. InformationSystems Journal.

Stieglitz(2014). Social Media Analytics: An Interdisciplinary Approach and its implication for Information systems. Business & Information Systems Engineering.

Arnes (1993). Retrieving and integrating data from multiple information sources. International Journal of intelligent and Cooperative Information Systems.

Watson & Wixom (2006). Real-time Business Intelligence: Best Practices at Continent Airlines. EDPACS the EDP audit, control and security newsletter.

Muntean, Mihaela and Muntean, Cornelia and Cabau, Liviu Gabriel (2003). Evaluat-ing Business Intelligence Initiatives With Respect To BI Governance. Munich Personal RePEc Archive.

Sousa (2008). Business and Information Systems Misalignment Model (BISMAM): an Holistic Model Leveraged on Misalignment and Medical Sciences Approaches. Jour-nal of Information Technology.

Horakh, Baars, Kemper (2008). Mastering Business Intelligence Complexity – A Service-Based Approach as a Prerequisite for BI Governance. Association for Information Systems.

Weil & Ross (2004). IT Governance: How Top performers manage IT Decision Rights for Superior Result. International Journal of Electronic Government Research, 1(4), 63-67, October-December 2005 63

Iyer & Raman (2011). Intelligent Analytics: Integrating Business Intelligence and Web Analytics. International Journal of Business Intelligence Research, 2(1), 31-45, January-March 2011.

Hevner, Park & March (2004). Design Science in Information System Research, Management Information Systems Quarterly

D. J. Power, C. Heavin, J. McDermott & M. Daly (2018). Defining business analyt-ics: an empirical approach. Journal of Business Analytics, 1:1, 40-53.

Allen Chamberlain (2003). Governance Frameworks to Manage Big Data as an Asset, EMC Corporation.

Asmaa Abduldaem, Andy Gravell (2019). Principles for the design and development of dashboards: literature review proceedings of intcess 2019- 6th international confer-ence on education and social sciences, 4-6 february 2019- Dubai.

Graeme Shanks, Rajeev Sharma (2011). Creating value from business analytics sys-tems: the impact of strategy. 15th Pacific Asia Conference on Information Systems: Quality Research in Pacific, PACIS 2011 (pp. 1-12).

Dave Chaff & Mark Patron (2012). From web analytics to digitalMarketing optimization: Increasing the commercial value of digital analytics. Journal of Direct, Data and Digital Marketing Practice (2012) 14, 30– 45.

Idri, Khoshgoftaar & Abran (2002). Investigation soft computing in case-based reasoning for software cost estimation. Engineering Intelligent Systems.

Ola El-Telbany, Ahmed Elragal (2014). Business-Information Systems Strategies: A Focus on Misalignment. Procedia Technology 16 (2014) 250 – 262

Aalst ,Reijers, Weijters, Dongen (2007). Business process mining: An industrial ap-plication. Information Systems 32 (2007) 713–732.

Katal, Wazid, Goudar (2013). Big Data: Issues, Challenges, Tools and Good Practices. Institute of Electrical and Electronics Engineers.

Constantiou & Jannis Kallinikos (2015). New games, new rules: big data and the changing context of strategy. 2015 JIT Palgrave Macmillan.

Wamba & AkterHow (2014). 'Big data' can make big impact: Findings from a systematic review and longitudinal case study. International Journal of Production Economics.

Fedorowicz J, Gelinas Jr UJ. Adoption and usage patterns of CobiT: results from a survey of CobiT purchasers. IS Audit Control J 1998;VI:45–51

Bin-Abbas & Bakry (2014). Assessment of IT governance in organizations: A simple integrated approach. we Volume 32, March 2014, Pages 261-267

B. Tuttle, S.D. Vandervelde / International Journal of Accounting Information Systems 8 (2007) 240–263

G. Hardy: Using IT governance and COBIT to deliver value with IT and respond to legal, regulatory and compliance challenges, Information Security Technical Report, vol. 11, no. 1, 2006, pp. 55–61.

Tallon, P. P. (2013). Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. Computer, 6(46),32-38.

Muntean, M., Muntean, C., & Cabau, L. G. (2013). Evaluating Business Intelli-gence Initiatives with Respect to BI Governance. Proceedings of the IE 2013 International Conference (pp. 1-6). MPRA.

ISACA, “COBIT 5: A Business Framework for the Governance and Management of Enterprise IT,” 2012.

Cabukovski V., Tusevski V. (2011) Aligning COBIT and ITIL with an IT Academic Courses. In: Wang Y. (eds) Education Management, Education Theory and Education Application. Advances in Intelligent and Soft Computing, vol 109. Springer, Berlin, Heidelberg

Luftman, J. N. (2003). Competing in the information age: Align in the sand (2nd ed.). New York: Oxford University Press.

Clementi & Carvalho (2006). Methodology for IT governance Assessment and design. Project E-Society: Building Bricks: 6TH IFIP International Conference on e-Commerce, e-Business, and e-Government (13E 2006), October 11–13, 2006, Turku, Finland (pp.189-202)

Sia &So (2007).An assessment of the package-organisation misalignment: Institution-al and ontological structures. European Journal of Information Systems p 568-583.

Appelbaum & Vasarhelyi (2017). Big data and Analytics in the modern audit en-gagement. Auditing A Journal of Practice & Theory p.36.

Lavalle, Lesser (2011). Big Data, Analytics and the path from insight to Value. MIT Sloan Management Review, p21-32.

Marinos, G. (2004). We're Not Doing What?; The Top 10 Corporate Oversights In Data Governance. Information Management, 9(14), 62.Friedman, T 2006. Key Issues for Data Management and Integration, 2006, Gartner Research. ID Number: G00138812, March.

Cheong, L. K., & Chang, V. (2007). The need for data governance: A case study. Proceedings of the 18thAustralasian Conference on Information System (pp. 999-1010). Toowoomba: AISEL.

Davenport, T. H., Barth, P., & Bean, R. (2015). How 'Big Data' Is Different. MIT SLOAN MANAGEMENT REVIEW,Spring(Special Collection: "Winning With Da-ta"), 7-9.

Akter, S., Gunasekeran, A., Wamba, S. F., & Dubey, R. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? International Journal of Production Economics(182), 113-131.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact.MIS Quarterly, 4(36), 1165-1188.

Cheong, L. K., & Chang, V. (2007). The need for data governance: A case study. Proceedings of the 18thAustralasian Conference on Information System (pp. 999-1010). Toowoomba: AISEL.

De Haes, S., & Van Grembergen, W. (2005). IT Governance Structures, Processes and Relational Mechanisms: Achieving IT/Business Alignment in a Major Belgian Finan-cial Group. Proceedings of the 38th Hawaii International Conference on System Sci-ences (pp. 237-247). IEEE.

De Haes, S., & Van Grembergen, W. (2008). Analysing the Relationship Between IT Governance and Business/IT Alignment Maturity. Proceedings of the 41st Hawaii International Conference on System Sciences (pp.1-10). IEEE.

Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assess-ment in the Big Data Era. Data

Science Journal, 2(14), 1-10.

Fernández-González, J. (2008). Business Intelligence Governance, Closing the IT/Business Gap. UPGRADE, 1(9), 23-30.

Luftman, J. (2004). Assessing Business-IT Alignment Maturity. Communications of AIS, 4(14), 99-128.

Van der Aalst, W. (2016). Process Mining: data science in action (2e ed.). Eindhoven: Springer.

Watson, H., & Wixom, B. (2007). The Current State of Business Intelligence. Com-puter, 96-99.

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data Quality in Context. Communication of the ACM, 5(40), 103-110.

Provost, F., & Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and

Data-analytic thinking. O'reilly media Inc.

Wu, S. P.-J., Straub, D. W., & Liang, T.-P. (2015). How Information Technology Governance Mechanisms and Strategic Alignment Influence Organizational Performance: Insights From a Matched Survey of Business and IT Managers. MIS Quarterly, 2(39), 497-518.

Wu & Zhang (2003). Machine learning and software engineering. Software Quality Journal, 2(11), 87-119.

Dai, W., Wardlaw, I., Cui, Y., Mehdi, K., Li, Y., & Long, J. (2016). Data profiling technology of data governance regarding big data: Review and rethinking. Information Technology: New Generations, 439-450.

Liao, S.-h. (2003). Knowledge management technologies and applications—literature review from 1995 to 2002. Expert Systems with Applications(25), 155–164.

Labrinidis & Jagadish 2012. Challenges and Opportunities with Big Data. Proceedings of the VLDB Endowment.

Negash, S. (2004). Business intelligence. Communications of Association for Information Systems, 13, 177- 195.

Olszak, C., & Ziemba, E. (2012). Critical Success Factors for Implementing Business Intelligence Systems in Small and Medium Enterprises on the Example of Upper Silesia, Poland. Interdisciplinary Journal of Information, Knowledge, and Management(7), 129-150.

Kanter, R. M., Stein, B. A. and Jick, T. D. (1992) The Challenge of Organizational Change (New York: The Free Press).

Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big data: the next frontier for innovation, competition, and productivity. New York (NY): McKinsey Global Institute; 2011.

APPENDICES

Remember to remove the numbering in front of the headings of appendices

Appendix 1. Definition design science model

Two main types of research define the IS discipline: behavioural science and design science.

The purpose of behavioural science is to produce and certify theories that provide a clear picture of the organization behavioural trends (Hevner, March, Park, & Ram, 2004). On the other hand, the design-science research aims at improving the organization capabilities by creating artefacts that "*define ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of information technologies can be effectively and efficiently accomplished*" (Denning, X. 1997; Tsichritzis, X. 1998).

Appendix 2. Big data three tiers architecture

Data accessing and computing procedures

For the first tier, the challenges mainly remains in the fact that big data are likely to be stored in various locations and grow rapidly. Therefore, an efficient platform has to be designed to be able to provide large and scalable storages. For instance, regulars' algorithms need the data to be uploaded within the main memory. This induce that the data has to be moved across various location which is expensive in term of network communication. Nevertheless, the data mining process differs according to the scale of the data. For small data endeavors, a personal computer equipped with a CPU is enough to mine the data. Concerning large data set (medium data) that may be spread over different locations and cannot fit into a single main memory, the proposed solution consist in using data mining platforms. In this case the data mining platform is going to parallel computing (Shafer et al. 1996) or collective mining (Chen et al. 2004) in order to aggregate differents data with several sources using *parallel computing programming*. In the case of Big Data mining, the size of the data is such that it impossible to deal with it using a single personal computer. Therefore, the big data processing framework uses a cluster of highly performing computers enabled by parallel programming like MapReduce. Software like MapReduce for instance, are used to ensure that in order to enable the perfect match for a query running a database with billions of entries, each data mining tasks will be efficiently divided into many ran on one or several computer (computing nodes). A computing node represents a data processing point that is part of a larger network.

Data privacy and domain knowledge

Tier II embodies the big data semantic and knowledge application. This tier refers to topics related to big data rules and regulation as well as domain information. Here, the two biggest challenges are 1) application knowledge and 2) privacy and data sharing. The first issues focus on knowledge discovery process and outcome whereas the second focuses on data maintenance, sharing and access. From a business perspective, application knowledge enables to use the correct data structure to model the data mining tasks so it provides an added value for the business (Kopanas et al. 2002). Here the challenge mostly stands in the fact that data miners are sometimes not aware enough of the domain knowledge regarding the data they are handling. Therefore, they are not able to grasp the benefices of data mining either from a business or scientific perspective. Concerning data sharing regulation, the first approach consists in limiting access to data by controlling authorized users. The second approach is to "anonymize" the data so that private information cannot be connected to a person personal records (Cormode and Srivastava 2009). The principal advantage of data anonymization is that it permits to almost freely share data across parties without limiting the access.

Big data mining algorithms

Finally, the last tier is big data mining algorithms. As described above, one of the main characteristics of big data is that each source of data is both autonomous and decentralized. Therefore, aggregating data from a different data source to a unique centralized hub in order to mine it is not the best option from a transmission cost perspective but also for data privacy issues. The only real option that remains is to develop the mining tasks in each distributed location. Here the only concern is the lack of overall view on the entire data set. Indeed, each location is going to have a precise understanding of the data mined in its site but when gather all the data mined from each site, it can become complex to reconstruct the puzzle. This phenomenon leads to a biased data set and to sub-efficient decisions. In order to cope with this challenge big data mining structures must create a systems that permit to efficiently exchange and integrate the data. For this purpose, model mining and correlation are critical methods that patterns that have been found out various data sources will be integrated to provide the overall mining goals. Model mining is a set of data, and more importantly of patterns that can be used in order to predict the behavior and pattern of a new set of data. The global mining process can be defined with two stages, first local mining and then global correlation. This process apply at the knowledge, data and model level. Concerning the data level, each location computes the statistics coming from its local data and exchange it between sites in order to create a global data view. At the model or pattern level, each site conducts site specific model mining and again share them with the other locations which potentially create a new model that summarize each mined model by aggregating patterns across all location (Wu and Zhang 2003). Finally at the knowledge level, correlating the patterns from each site permits to analyze the relationship, similarities and differences between each data source. This will, thereafter, permit to ensure that the decisions will be taken based on the awareness of each local data and model as well as on the global pattern. In other words, the challenges for the organization is to develop the capability to learn locally and then efficiently merge the multiple source of information at the organization level in order to generate a general overview of the big data behavior.

The second big challenge stands as the organization ability to mine sparse, incomplete and uncertain big data. The fact that Big Data can be sparse can impede the miners to draw reliable conclusion as the number of data point is to low. This issue comes from the high number of data dimension that big data normally has. Data dimension refers to the number of attributes a data set/sets can have. The bigger the data dimensionality is the harder it gets to draw conclusion and to identify trend from the data. Likewise, from a data mining algorithm perspective a high dimensionality impacts both the reliability of the model and increase the difficulty to draw conclusions. Here a common approach is to use attribute selection or to diminish the number of dimensions (Wu et al, 2012). Concerning uncertain big data, each data field is potentially exposed to distribution errors. The sources of data uncertainty come from domain application's lack of ability to collect and interpret the data. For instance, the data generated from GPS can be uncertain because of the limited accuracy of the technology or the income communicated from a person is a range [40-80k per year]. This impedes regular data algorithms to process the data. One solution for tackling this issue is to embed the data distribution into the parameters of the model. Error data mining uses both the mean and variance for each data field in order to build a "Naïve Bayes" clustering model. Finally for incomplete data the most common approach is to create a learning model to forecast the possible values for each data item missing basing the analysis on the formerly observed data.

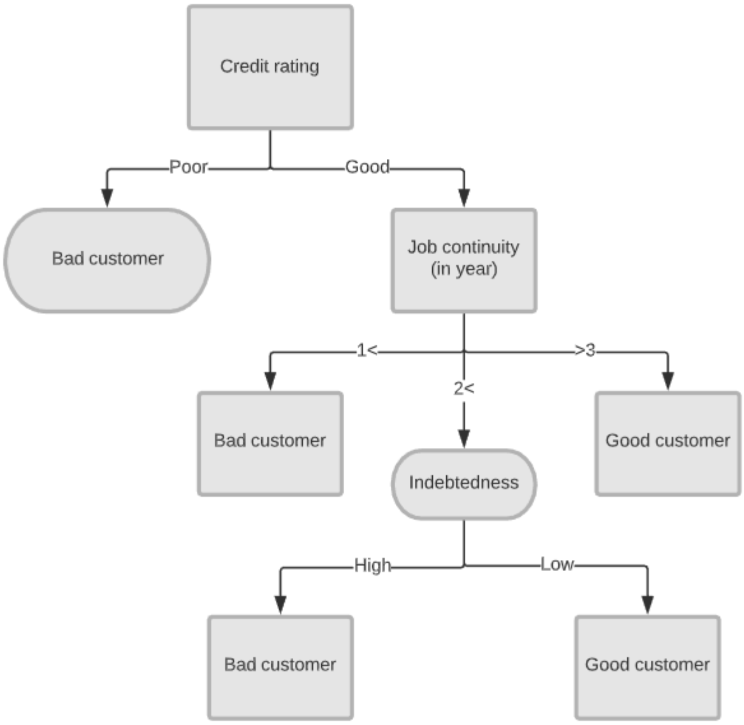
In a third time, the last big challenge for algorithm bid data mining concerns complex and dynamic data. The rise of the social media, worldwide servers and internet backbones has driven the increase of changing and dynamic data. But the volume of the data isn't the only factor impacting the organization ability to perform its algorithms in given set of data. Indeed the nature of the data is also changing as the data becomes more and more complex. Even if the complex unstructured data present on the social media for instance have a high potential for the organization, using this complex data is a critical challenge in big data mining. In the big data context, it is huge challenge to create semantic model to fill the semantic gap with the very different heterogeneous data sources.

Appendix 3. Objectives of process mining

The process mining field finds its foundations in the information technologies like Enterprise Resource Systems, Customer Relationship Management systems, Supply Chain Management systems and others B2B systems. Indeed those management systems business event called "event logs". Process mining methods use event logs for process discovery, verify compliance, identify flaws and differentiate the process paths (Van Der Aalst, 2016). An event log is composed of different features, in the first place each event correspond to an activity. Moreover, each event refers to a case and may have one or several performers (Figure X). Finally each event detains ordered timestamps. From organizational perspective, the data related to the performer permit the company to increase its understanding concerning the different involvement and relationship of the employees that make use of systems (who uses the process). The objectives is to cluster human resources by roles or to display the relationship between performers. The case perspective on the other hand, indicates the characteristic of a certain case. While using cases, the focus is mostly oriented to identifying the different process path or the performer of the case. Also, cases can be used to identify the values that refers to the event (Van Der Aalst, 2007). Finally the activity ID and the timestamp permit to identify and situate a certain case. Process mining as described above answers a need from organization to monitor the entity processes. As an example new legislation like Sarbanes-Oxley (SOX) consistently oblige corporate to improve their ability to monitor and control their business processes. However, beyond the organization's need for compliance with the law, Business process mining also permit to monitor the company's performance. In short, business process mining embodies the discipline of automatically creating models that describe an event log and provide the company with an insight on what are its current processes (case), how are they managed (process path), by whom (performers) and when (timestamp).

Appendix 4. Definition of the machine learning techniques

Rule induction:
The rule induction develop a decision tree from a training data set. Based on the training data set, rules that are going to be applied for the further data analysis are developed and used to gather useful information from large set of data of unstructured data (Bose, 2001). In the case of a bank that tries to identify the customer profile of client that don't reimburse their loan, the rule induction method would use a data set including the records of customers that were formerly granted loans. This process would consist in selecting the attributes and then to split this attributes into different categories, in our case the first attribute could be "credit rating" split into good or bad and identify to which group belongs the delinquent customer. The credit rating "bad" directly identifies the group as a risky customer but the group credit rating good is mixed (composed both with good and bad customers) therefore leading to the use of other attributes like job continuity, Indebtedness etc… The decision tree is added attributes until the threshold of bad customer contained in a group remains acceptable. Here it is also important to emphasize that an overfitting training data set diminishes the accuracy of the decision tree over new set of data.
Figure A: Process mining scheme



Neural Network:

A Neural Network is sequence of input that recognizes connections in a set of data. The main characteristic of this discipline is that it mimics the way a human brain works, it optimize the output by adapting to the change of the inputs. Neural networks embody the foundation for artificial intelligence. In practice this network is composed with an input layer, a hidden layer, an output layer and a function between each hidden layers. Every connection is given a weight that will change the incoming signal and thereafter pass it to another node. Therefore the signal of a node is not more or not less than the sum of the weighted signals sent to this node (Bose, 2001). The advantage of the Neural Network is that it can comprehend any classification functions.

Case-based reasoning:

The case-based reasoning is the method using solution found for a problem to solve a similar new problems. This discipline consist in storing the output of formerly resolved problem and to reuse them for machine learning. In this case both problem and solution are stored. Based on the new problem the algorithm tries to find a match the former solution to propose a new solution for the new problem. The quality of a case-base reasoning outcome depends on the diversity and quantity of the former solutions and problem statements. Furthermore, the analysis can be divided into four steps: collecting the most similar cases, adapt the data retrieved from the selected cases, optimize the proposed solution and store the solution for future experiment (Idri, 2002).

Even though machine learning unify those techniques into a single domain, the aforementioned methods behave in very different according to the data structure. The case of inductive reasoning the method will fairly handle large, noisy and incomplete set of data (Idri, 2002). Moreover, it will permit to predict accurately the behavior of new set of data. On the other hand, Neural Network will handle noisy and incomplete data but the method is very time consuming and large set of data often limit the use of the method. Moreover, when applying those method to business their use is also very different. Rule induction will be used in the financial sector mostly for predictive analysis and classification but also in most of the sector using machine learning like Marketing or Web Analysis. Neural network will mostly be used to predict financial trends and in some cases provide software cost estimation.

Appendix 5: Challenges and critical factor of success of Web analytics

According to the study of Dave Chaffey (2012): *"Many companies are failing to utilize core business analytics best practices and are therefore not getting the potential return from web analytics that they could".* One of the biggest reason for these failures apart from the lack of resources and capabilities is the lack of use of fudamental web analytics methods as customer journey analysis or A/B testing. The customer journey analysis allow the company to map the customer online purchasing process in order to produce useful consuption patterns. This type of analysis can be both desciptive or predictive (Bain&Company, 2018). This discipline very usefull to reap the benefice of complex data set consist in anlysising all the customer interaction with the online plateform starting when she/he enter the web site. Moreover, it gathers all the data relative to the customer behavior and link the data gathered in diferent department owning certain part of the process in order to provide a complete picture on how the customer experience its journey. Finally it enlighten all the disceprency of the customer journey including customer's waste of time.

Neverthless, comapnies using web analytics are now coping with new challenges. The biggest barrier to web analytic performance in the first decade of the twenty first century was technologie and the lack of capabilities integrating information technologies. However, the trend is now evolving and companies increasigly point out to the factors people and process as their major impediment to perform in web analytics. Dave Chaffey (2012) uses the convertion rate to define the correct inetgration of web analytics in the company. The convertion rate define the ratio of potential customer that take a decision in the company interest (buying, publishing, etc...). According to this study the four factors that were the most closely related to improving the web site conversition rate were either based on web analytics process integration or on the human capabilites of the company:

1. *"Perceived control over conversion rates."*
2. *"Having a structured approach to CRO."*
3. *"Having someone directly responsible for CRO."*
4. *"Incentivizing staff based on conversion rates."*

Concerning the major challenges, the top two issues are related to the companies' capability of efficiently using data provided throughout web analytics tools to improve the conversion rate. Here the study shows that web analytics capabilities in numerous cases are used almost exclusively for supporting the business (reporting, running the current project). This finding clearly support the fact that not enough human capabilities are dedicated in finding developing the business but more on supporting it.

The aforementioned challenges shows the companies' need for benchmarking the capabilities of web analytics. Hamel (2009) proposes a model for improving the use of web analytics that is based on six core capabilities:

1. *Management, Governance and Adoption: Consist in* creating a clear governance and management framework that define responsibilities.
2. *Objective definition:* ensure the goals of web analytics are correctly defined
3. *Scoping:* Define strategy based on either conversion rate optimization or digital marketing.
4. *Analytics and expertise:* Assess of the current human analytics capabilities over entire the company
5. *Continuous improvement process and methodology:* Creation of an auditing process based on performance and metrics. Shift the role of the analytic expert from reporting to a performance driven.
6. *Tools technology and data integration:* Ensure that the company owns all the technical capabilities.

Appendix 6: A framework for social media analytics

In its framework aiming at increasing the efficiency of social media analytics for the organization by showing how the discipline impact the organisation, Kurniawati (2013) describes three main concepts: *”organisational motivation”*, *”social media capabilities”* and *”benefits”*. For each of these concepts the framework identifies factors that have been reflected as main drivers for success in implementing social media analytics. The organization motivation factor embodies the organization ability to lean toward the succesful vision of handeling the discipline. The most important motivational factors driving success is the organisation will to develop *”an in-depth understanding of the customer values, preferences, behaviours”* Kurniawati (2013). Moreover, *”gathering ideas about brand and product”*, *”determining the impact of online campaings”* and *”indentify social inflencers”* also are shown as being critical motivational factors for success. Motivations lead the organisation to create social media capabilites whose goal is to gather knowledge about online context and content as well as the business implication of these information. The framework indentifies *”sentiment analysis”* as being the capability the most correlated to the organisation success. This analysis identify the sentiment or the behavior of the population concerning a particular topic. Moreover, *”insight mining”* that aim at determing an insight into the population behaviour, will and preferences as well as *”trend analysis”* and *”competitive analysis”* are also capabilities that have been shown as leading the organisation to sucess.

Table B: Kurniawati, Shanks, Bekmamedova (2013).

Dimensions	Capabilities	Description
Motivation	*”Provide customer insights.”*	*”The need to have an in-depth understanding of customer values, preferences, behaviours.”*
	”Gather ideas about brands, products and services.”	*”The need to gather new ideas about brands, products and services, including online feedback.”*
	”Determine the impact of online campaigns”	*”The need to measure the return on investment (ROI) and effectiveness of online marketing and outreach initiatives.”*
	Sentiment analysis	*”Determine the sentiment polarity (positive, negative or neutral) or attitudes to a particular issue.”*
	Insight mining	*”Discover insight into customer behaviours, intentions, and preferences.”*

Social media Analytics capabilities	*Emerging issue and trend analysis*	*Track and monitor issues and how they change over time.*
	Influence analysis	*Identify the key people or communities that have made significant contributions to a particular issue*
	Competitive analysis	*Track and monitor comments about brands and products of competitors*
Benefits	*Marketing strategy improvement*	*Create and refine marketing strategies, initiatives and channels in order to effectively deliver messages to targeted customers*
	Better customer engagement	*Provide two ways of communication with targeted customers, based on their values and preferred channels*
	Customer service improvement	*Provide timely and appropriate responses to customer feedback*
	Better brand awareness and reputation management	*Monitor and maintain brand and product reputation in the market.*

successfully grasp the benefices of a package. A deep structure misfits is characterized by the absence of a *"real-world things, properties, states and transformation are missing or incorrectly represented in the system"* (Sia & Soh, 2007, p. 572). The surface misfit happends when the users misuse the original ERP inteface. Therefore there is a gap between the way the ERP was suposed to be used and the actual way that users handle the data. These misaliments are generaly not impacting the core of the business and relate to a matter of conveniance of the end-users. By combining the two approaches deep/surface and impose/volontary structure Sia & Soh (2007) define four categories of misfits: *"imposed-deep misfits, imposed-surface misfits, voluntary-deep misfits and voluntary-surface misfits" ;* and cluster them into three categories relative using a software application perspective: data, process, output . A data misalignment regards the lack of capability the organization has to capture the information into the schema of the ERP solution. A data misalignment is always considered as deep structure. Also, a process misalignment is related to the lack of capability the organization has to integrate is internal process or functions with the packages attributes and pre-established procedures. The output misfit relates to the lack of integration between the package and the company needs in term of data visualization. The output misalignment is always a surface structural misalignment. Also Strong (2010) proposes three additional misalignments that arise from the creation a latent structure. The latent structure represent the second structure that is created due to surface or deep structure misalignment between the functionalities of the ERP and the way the employees use those functionalities.

Appendix 7. Resource-based theory

This research bases its foundation on the resource-based view of (Barney 1991, Wade & Hulland 2004) that emphasize that organizational ressources is the most important factor in improving the firm performance. The company's ressource can therefore be tangible or intangible and includes resources that are both technical and human.

The resource-based view also divides the different resources of the company into four categories that compose the VRIN framework: Valuable, Rare, Inimitable and Non-substituable (Amit and Schoemaker, 1993; Barney et al., 2001). In the first place, the criterion *''valuable''* allow the organization to improve its revenues and to decrease its costs. In the second place, the *''rare''* resource emphasize the fact that the asset is not commonly used within the market and therefore represent an almost unique competitive advantage for the company. In the third place, an *''inimitable''* resource defines an asset that other organization can hardly copy or. Finally, a *''non-substiable''* defines the impossibility for other firm to produce the same capability from any of their resources. Former literature also highlight the importance of the organizational dimension that focus on managing the four aforedmentioned factors in order to fully leverage their competitive advantage. Resources and capabilities compose the main foundation for the Ressource Based Therory. On one hand, resource define the asset itself and capabilites are the subset of competencies and skills that are unique for the firm and improve the efficiency of the core resources of the organization(Makadok, 1999). The Resource Based Theory emphasize that the competitivity of an organization remains on its ability to steer its core resources to develop a competitive advantage (Grant, 2002). BA represent one of the most critical capability for the firm to achieve organizational performance as well as a competitive advantage in the current big data environment. This is why, the four dimension of the VRIN for BA resource may represent the foundation for creating a superior firm performance (Akter, S ,2016). Nevertheless it appears that few research on BA have used this framework in order to define the capability needed to effienciently predict the performance a organizations (Abbasi et al., 2016; Phillips-Wren et al., 2015). Therefore it has been argued that the organization ability to perfom in a data oriented environment can be improved only when the firm's capabilities are valuable, rare, inimitable and non-substituable (Akter ,2016).

Appendix 8. Misalignment concept and semantic a Business and IS approach

As this approach of organizational misalignment is based on the medical science foundation, we define below the misalignment semantic that embodies the concept (Table C). This concept adaptation is based on the observation of criteria that correlate medical science with the organizational science. Indeed, both organizational and medical science study complex systems that require a numerous intra-system entities to function together. For instance, the organization architecture components must be designed to function together just like the human's organs (Sousa, 2008).

Table C: Misalignment - semantics (Sousa, 2008)

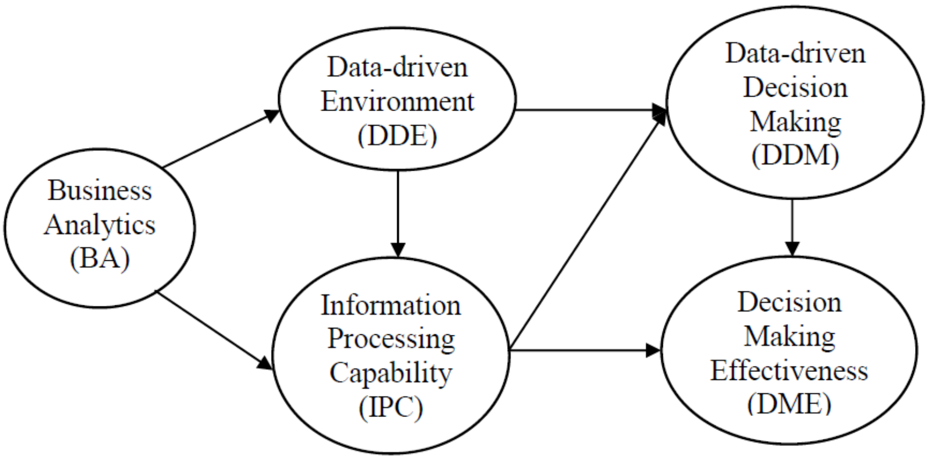
Classification	Organ system	Defines the architectural component involved in the system.
	Symptom	The subjective proof of misalignment that impact the entity employees.
	Sign	The objective proof of misalignment that can be observed both internally and externally.
	Syndrome	Represent a set of both symptoms and sign
Management	Etiology	The root cause of the misalignment
	Diagnosis	The procedure of investigating the misalignment in by observing signs and symptoms using a research method (test, questionnaire).
	Therapy	Set of actions taken in order to solve the misalignment identified throughout diagnosis.
	Prophylaxis	Governance model, framework, guidelines and set of principles aiming at preventing the misalignment.

Appendix 9. Consequences of actual misalignment throughout the layers of the company (Van Groenendaal, 2015).

ERP Misalignments	Consequences	Layer
Poor usability by target community	• Generating incorrect data • Loss of trust in the ERP system by end users	Operational
Complexity and poor visibility of ERP calculation logic	• Decision-making with questionable output • Lack of control about data output • Loss of trust in the ERP system by end-users	Operational
Incompatible input data	• Creating wrong input • Creating complex reports • Creating incorrect data • Wrong use of input data	Operational
Incompatible terms and meanings	• Loss of trust of the ERP system by end-users	Operational

	• Creating wrong input • Wrong use of input data	
Missing non-transactional functionalities	• Missing validation function Operational • Generating incorrect data • No control over unauthorized people •Higher risk of fraud	Operational
Poor system quality and performance	• Loss of trust in the ERP system by end-users	Operational
Incompatible IT infrastructure	• Loss of trust in the ERP system by end-users Tactical • Need for change in the IT infrastructure	Tactical
Poor output quality or accuracy	• Wrong decision-making due to poor quality/accuracy of data	Tactical
The complexity of reports and interface	• Wrong decision-making due to complex data	Tactical
Poor data visibility	•Uncompleted data reports • Use of extra inventory	Tactical
Poor data accuracy	• Incorrect information processing • Incorrect reports	Tactical
Inappropriate data presentation and output format	• Incorrect reports Tactical • Inflexible reports	Tactical
Conflicts with management philosophy and organizational structure	Need for change in the business processes • Need for change in the organizational structure	Strategic

Appendix 10. Research model business analytics capabilities process view (Cao, 2015)



Appendix 11. Expert data interview

Program manager 1 Interview:

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	7	4
	Web Analytics	5	2
	Statistical analysis	8	8
	Forecasting	9	5
	Query and analysis	8	8
	Machine Learning	5	3
	Simulation & scenario development	10	7
	Dashboards	8	6

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	8	6
	Understand the customer	10	8
	Make real time decisions	9	7

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	8	8
	BA capability of integrating data	8	6
	BA capability of analyzing data	10	7

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	9	8
	Policies and rules guiding BA	7	4
	Department structure adapted to BA	10	7
	BA integration into business process	8	8
Strategical: Data-driven environment	Data-driven decision for process creation	9	7
	Data availability for taking decision	10	7

Program manager 2 Interview:

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	9	9
	Web Analytics	8	7
	Statistical analysis	8	7
	Forecasting	10	8
	Query and analysis	10	8
	Machine Learning	8	7
	Simulation & scenario development	9	7
	Dashboards	9	6

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	9	7
	Understand the customer	10	8
	Make real time decisions	9	7

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	9	7
	BA capability of integrating data	9	8
	BA capability of analyzing data	9	7

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	9	9
	Policies and rules guiding BA	8	6
	Department structure adapted to BA	8	7
	BA integration into business process	9	9
Strategical: Data-driven environment	Data-driven decision for process creation	10	6
	Data availability for taking decision	10	8

Program manager 3 Interview :

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	9	8
	Web Analytics	5	5
	Statistical analysis	8	7
	Forecasting	10	7
	Query and analysis	10	8
	Machine Learning	8	7
	Simulation & scenario development	9	6
	Dashboards	9	8

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	9	8
	Understand the customer	10	8
	Make real time decisions	9	8

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	10	6
	BA capability of integrating data	10	8
	BA capability of analyzing data	10	8

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	9	9
	Policies and rules guiding BA	8	9
	Department structure adapted to BA	8	9
	BA integration into business process	10	5
Strategical: Data-driven environment	Data-driven decision for process creation	10	6
	Data availability for taking decision	10	8

Business Analyst 1 interview:

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	1	1
	Web Analytics	1	1
	Statistical analysis	7	7
	Forecasting	8	8
	Query and analysis	8	8
	Machine Learning	9	9
	Simulation & scenario development	4	4
	Dashboards	10	10

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	10	10
	Understand the customer	10	10
	Make real time decisions	8	8

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	9	9
	BA capability of integrating data	9	6
	BA capability of analyzing data	9	9

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	9	9
	Policies and rules guiding BA	8	8
	Department structure adapted to BA	6	6
	BA integration into business process	7	7
Strategical: Data-driven environment	Data-driven decision for process creation	10	10
	Data availability for taking decision	8	8

Business Analyst 2 interview:

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	8	8
	Web Analytics	4	3
	Statistical analysis	8	7
	Forecasting	7	5
	Query and analysis	9	6
	Machine Learning	7	2
	Simulation & scenario development	8	8
	Dashboards	8	8

Part 2: Business Analytics at the tactical level

-

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	10	6
	Understand the customer	9	7
	Make real time decisions	8	4

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	9	8
	BA capability of integrating data	8	8
	BA capability of analyzing data	9	10

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	9	8
	Policies and rules guiding BA	7	7
	Department structure adapted to BA	7	7
	BA integration into business process	8	8
Strategical: Data-driven environment	Data-driven decision for process creation	8	9
	Data availability for taking decision	8	9

Data Analyst 1 interview:

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	10	9
	Web Analytics	5	3
	Statistical analysis	10	7
	Forecasting	9	7
	Query and analysis	10	8
	Machine Learning	9	5
	Simulation & scenario development	8	6
	Dashboards	8	7

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	9	5
	Understand the customer	10	7
	Make real time decisions	10	4

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	10	9
	BA capability of integrating data	10	8
	BA capability of analyzing data	8	7

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	7	9
	Policies and rules guiding BA	5	5
	Department structure adapted to BA	7	6
	BA integration into business process	9	9
Strategical: Data-driven environment	Data-driven decision for process creation	9	9
	Data availability for taking decision	10	10

<u>Data Analyst 2 interview:</u>

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	6	5
	Web Analytics	6	2
	Statistical analysis	10	7
	Forecasting	9	9
	Query and analysis	10	10
	Machine Learning	9	5
	Simulation & scenario development	10	7
	Dashboards	10	7

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	10	6
	Understand the customer	10	8
	Make real time decisions	10	7

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	10	9
	BA capability of integrating data	10	9
	BA capability of analyzing data	10	9

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	10	9
	Policies and rules guiding BA	8	5
	Department structure adapted to BA	9	7
	BA integration into business process	9	9
Strategical: Data-driven environment	Data-driven decision for process creation	10	7
	Data availability for taking decision	10	8

Data Analyst 3 interview:

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Operational: Business Analytics tools	Data and text mining	8	5
	Web Analytics	7	4
	Statistical analysis	10	7
	Forecasting	9	9
	Query and analysis	10	10
	Machine Learning	9	5
	Simulation & scenario development	10	7
	Dashboards	10	7

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	10	6
	Understand the customer	10	6
	Make real time decisions	10	7

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	10	7
	BA capability of integrating data	10	9
	BA capability of analyzing data	10	10

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the component for the department (1-10)	Performance of the component within the department (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	7	7
	Policies and rules guiding BA	7	3
	Department structure adapted to BA	7	7
	BA integration into business process	7	7
Strategical: Data-driven environment	Data-driven decision for process creation	7	5
	Data availability for taking decision	8	5

<u>Data Analyst interview:</u>

Part 1: Business Analytics at the operational level

Construct	Component	Importance of the indicator for the Department (1-10)	Presence of the component within the Department (1-10)
Operational: Business Analytics tools	Data mining	8	7
	Web Analytics	3	1
	Statistical analysis	8	5
	Forecasting	10	8
	Query and analysis	10	10
	Machine Learning	8	6
	Simulation & scenario development	10	9
	Dashboards	10	8

Part 2: Business Analytics at the tactical level

Construct	Component	Importance of the indicator for the organization (1-10)	Presence of the component within the organization (1-10)
Tactical: Decision making effectiveness	Ability to quickly respond to change	10	8
	Understand the customer	10	9
	Make real time decisions	10	10

Part 3: Business Analytics at the analytical level

Construct	Component	Importance of the indicator for the organization (1-10)	Presence of the component within the organization (1-10)
Analytical: Information Processing Capabilities	BA capability of capturing data	10	10
	BA capability of integrating data	10	8
	BA capability of analyzing data	10	10

Part 4: Business Analytics at the strategical level

Layer	Component	Importance of the indicator for the organization (1-10)	Presence of the component within the organization (1-10)
Strategical: Data-Driven Decision making	Department strategy guiding BA	10	9
	Policies and rules guiding BA	7	5
	Department structure adapted to BA	8	8
	BA integration into business process	9	7
Strategical: Data-driven environment	Data-driven decision for process creation	10	9
	Data availability for taking decision	10	8