UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Innovation and Industrial Management

Big Data-driven Innovation:

The role of big data in new product development

Deni Redzepovic and Theodoros Peristeris

Supervisor: Erwin Hofman
Master Degree Project No. 2015:45
Graduate School

BIG DATA-DRIVEN INNOVATION: THE ROLE OF BIG DATA IN NEW PRODUCT DEVELOPMENT

By Deni Redzepovic & Theodoros Peristeris

School of Business, Economics and Law, University of Gothenburg,
Vasagatan 1, P.O. Box 600, SE 40530 Gothenburg, Sweden

Contact: redzepovic.deni@gmail.com or peristeris.theodoros@gmail.com

Abstract

The outcome of the Big Data hype is a widespread delusion that applying Big Data will instinctively enhance and improve current business activities. Organizations have identified the benefits and disruptive potential of Big Data such as; value creation, business idea generation, and innovation. Yet, most organizations do not understand the key factors that go into a successful Big Data implementation or Big Data's role for specific objectives. Previous studies show that unsuccessful implementation of Big Data across organizations is immense, only 27% of the companies being investigated described their Big Data actions as successful (Capgemini, 2015). Nevertheless, exponentially increasing data streams enable novelty in methods and processes for New Product Development. Making Big Data an important enabling element of innovation and sustainability.

The purpose of this study is to explore and understand what role Big Data holds in the New Product Development and what factors influence a successful implementation in New Product Development processes. A qualitative investigation through multiple case studies of diverse companies in the Netherlands was executed to explore and compare key elements of Big Data in the context of innovation, and more specifically in the pre-development and formal development phases of New Product Development. In summary, the empirical findings show that the role of Big Data in the New Product Development is highly dependent on the ability to understand the specific objective or problem, and examine if using Big Data is the right approach for solving that problem. There is a prerequisite for securing distinct resources and organizational capabilities to succeed with implementing Big Data into the New Product Development. Other important factors that need to be well considered by organizations when forming an implementation strategy is organization's data maturity and effective change management, especially if the organization is utilizing more traditional innovation processes. However, novel methods rely heavily on extensive and varied data which translates in an adoption urgency to sustain competitive advantage and secure responsive innovation.

Keywords: *Big Data, Innovation, New Product Development, Business Analytics, Implementation,*

Acknowledgements

First and foremost, we wish to express our sincere appreciation to our supervisors, Evangelos Bourelos and Erwin Hofman, for providing us with their support through helpful guidance and encouragement during our master's thesis journey. We would also like to thank University of Twente for hosting us.

Further, we would like to express our gratitude to the Institute of Innovation and Entrepreneurship, Sten A. Olsson foundation, and Professor Rick Middel for making it possible for us to pursue with this research project. A special thanks to Professor Jan van den Ende at Rotterdam School of Management - Erasmus University for sharing his valuable time and providing helpful advice.

Lastly, we would like to thank all the participants who agreed to be interviewed and share their knowledge and experiences with us.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Background

Given modern industry dynamics, New Product Development (NPD) is becoming more and more important to achieve growth, sustainability and competitiveness. Several technologies and methods are used to obtain value and add them to products and services. In order for an organization to be able to compete and capture market share, it requires to find alternative ways of acquiring a competitive advantage. Despite of what many decision-makers believe, designing a sustainable model of value creation and innovation is more efficient in the long term than preserving the traditional approach (Nidumolu et al., 2009). An important enabling element of innovation and sustainability is the use of emerging technologies, and the constant improvement of value creation processes.

Organizations have to innovate in order to survive and evolve. To achieve that, it is vital to have access to new processes, technologies and knowledge. Therefore, firms should possess the necessary agility to leverage them and create the new business possibilities. Currently, some of the rewarding opportunities come from technological achievements and specifically, electronic forms of information.

In recent time, Big Data has gone from an unutilized phenomenon to a widely accepted concept to improve business and operational activities. Companies recognize the benefits and the disruptive potential of Big Data such as; value creation, business idea generation, and competitive advantage. However, the resources, capabilities, strategy, and knowledge base needed to fully embed Big Data in their operations might not exist within the company itself. As the technological ability to capture and store limitless streams of data has increased at a rigorous rate, the organizational capacities and resources to bundle, analyze and manage the vast quantities of unconnected data are lagging behind (Manyika et al., 2011; Parmar et al., 2014).

Big Data is a broad term generally referring to very large data collections that impose complications on current data processing tools for handling, analyzing and managing such. Ohlhorst (2013) states that a general definition can be abridged into: *"Big Data defines a situation in which data sets have grown to such enormous sizes that conventional information technologies can no longer effectively handle either the size of the data set or the scale and growth of the data set"* (Ohlhorst, 2013, p. 18).

The use of Big Data for New Product Development is providing the means to new opportunities. But is it so simple to acquire, analyze and finally implement the changes that Big Data suggest? According to a research conducted on IT-professionals in 18 countries worldwide, show that 60 percent of the participants consider that Big Data could improve competitiveness and management decisions, but only 28 percent of the companies believed that they are extracting strategic value from this process (European Commission, 2013).

There are many factors that go into successful Big Data implementation. Previous research on the topic shows that unsuccessful implementation of Big Data across organizations is immense, only 27% of the companies investigated by Capgemini described their Big Data actions as successful (Capgemini, 2015).

1.2 Purpose

The purpose of this research is to understand what role Big Data plays in the New Product Development and what factors influence on how companies can successfully organize and make use of Big Data in their New Product Development processes. The research will be explorative due to the novelty of the topic, where the purpose is to investigate how Big Data is used to identify potential innovation and support the pre-development, as well as the formal development phase of the New Product Development.

A theoretical framework will examine current organizational theories and methods for enabling innovation, business growth and adding value through Big Data. Additionally, the empirical section of this thesis will through a multiple case study go into depth with how Big Data is interpreted, processed, applied and deployed to assist the New Product Development process. By providing a foundation of understanding of what role Big Data plays in the New Product Development and which specific organizational drivers/factors exist in terms of embedding Big Data in innovational functions, companies will have a basis for where to focus their efforts.

1.3 Research Question

The objective of this research is to explore the role of Big Data in the context of innovation management and New Product Development. The research aims to identify how Big Data is currently managed and implemented for driving innovation and New Product Development. Further, what benefits and managerial challenges are connected to using Big Data within New Product Development. By providing a foundation of understanding, as to what specific key success factors exist in terms of stimulating the pre-development and the formal development phase - in the New Product Development - through the usage of Big Data, companies will potentially have a basis for where to focus their efforts.

The arisen Big Data movement has given strong incentives to conduct both corporate and academic research. Yet, existing literature was found to be inadequate regarding the usage of Big Data in the New Product Development. Theory was either focused on Big Data or New Product Development, but only a fraction of all subject-relevant publications is exhaustively investigating their integration. Specifically, great amount of literature investigated the emerging usefulness of Big Data for commercial purposes other than driving innovation.

Furthermore, many scholars were examining various aspects of New Product Development, while some of them were suggesting an expanded use of data. Despite the amount of research conducted in the field, finding theoretical frameworks of integrating Big Data analytics into the NPD process proved to be a significant challenge. For that reason, this exploratory research will be investigating the role of Big Data in the NPD process by combining theories from two different domains (NPD and Big Data) grounded on the empirical findings derived from diversified corporate settings. Thus, aiming to address and answer the main research question:

- ***What is the role of Big Data in the New Product Development?***

Together with the following sub-questions which will assist in answering the central research question:

- ***What are the benefits of using Big Data in the New Product Development?***
- ***What are the managerial challenges of using Big Data in the New Product Development?***

1.4 Empirical Setting

In recent time, the Netherlands has shown high degree of Big Data initiative and utilization from both the corporate sector and public sector. Many global and leading companies, which are headquartered in the Netherlands, acknowledge the importance of Big Data. Furthermore, it was evident that a lot of new Big Data startups were emerging within the Dutch boarders - offering solutions, consultancy or data-driven products. The Netherlands appeared to be upfront in the European Big Data development, and for that reason we primarily focused to conduct our research in that particular setting.

1.5 Limitations

The intention of the thesis is not to cover Big Data as an isolated phenomenon, as there is a notably large amount of research on the topic, but rather on the utilization and management of Big Data in New Product Development processes.

The study will not focus on the technical design or functionality of Big Data implementation and management, e.g. engineering and programming. The central perspective of this thesis will be through a management of technological innovation point of view. However, it is highly likely that key technical factors or attributes linked to the usage and implementation of Big Data, that exert influence, will be explored.

1.6 Disposition

This research will commence with a theoretical framework, which is divided into three parts. The first part will investigate applied concepts and models of the pre-development and the formal development stages of New Product Development. Chapter 2.1 (Innovation Management) will explore the main tools and processes being used for New Product Development (New concept development model and Stage-Gate model), as well as their implications to better comprehend the role of Big Data in applied processes. Furthermore, the chapter will explore more recent and modern methods that are being implemented which are interconnected with Big Data, such as the Lean startup method.

The second part of the theoretical framework will proceed with introducing Big Data grounded in contemporary literature on the topic. This part will cover relevant research theories, concepts and dimensions of Big Data in context of business and innovational activities. Chapter 2.2 (Big Data) will through existing literature on Big Data explore the relevant meaning of Big Data, as well as its drivers, benefits and challenges.

The third and last part of the theoretical framework will provide a theoretical investigation of Big Data implementation features and factors. Chapter 2.3 (Implementing Big Data) will examine the main factors that go into the implementation of Big Data, as well as draw connections to the key organizational capabilities that are required based on the existing theory.

Following the theoretical framework, this thesis will present the applied methodology for answering its central research question and sub-questions. This section will explain and outline the research strategy, research design, research methods and sampling method used to obtain the empirical findings.

Subsequently, the empirical findings will be presented in structured tables and followed by an analysis of the empirics. The last section of the thesis will include conclusions in conjunction with recommendations and implications of the role Big Data holds in the New Product Development.

2. Theoretical Framework

2.1 Innovation Management

Innovations are adopted when corporations combine and integrate them in purposive methods into current operational activities (Tuomi, 2002). To comprehend the management and integration of innovations there needs to be an understanding of concepts and definitions. First of all, it is essential to define innovation. However, a simple and general definition of innovation is almost impossible. Spontaneous answers to the question “what is innovation?” would often resemble a definition of the term “invention”. According to Dodgson et al. (2008), “invention” is the conceptualization of a new idea and its practical application, opposed to “innovation” which includes the entire process of commercializing a new idea. Chesbrough et al. (2006) define innovation as invention which is implemented and introduced to the market.

The ability to capture value from innovation is an essential part of achieving sustainable competitiveness. Grant (2010) argues that successful innovation management can generate value and long-term financial success. Additionally, it can enable companies to brand themselves as a productive, creative and stimulating corporate environment to attract and retain expertise and knowledge. On the other hand, if companies are unsuccessful in their management of innovation activities they have to confront critical issues such as negative results, and inability to retain and attract human resources (Grant, 2010).

Corporations have to possess the ability to adapt and evolve for ensuring survival in today’s global business sphere. The literature on the topic of business studies and innovation management highlight the importance of being able to adapt to the rapidly changing market conditions (Trott, 2011; Afuah, 2002; Dodgson et. al., 2008; Davila et al., 2013). Accordingly, market leadership is constantly altered after an introduction of new or improved product, meaning that innovations can disrupt the current market formations. Innovations can be incremental (i.e. enhancements or improvements in product design and new features) or discontinuous (i.e. radical or newly discovered technological innovation without an existing market).

The Scope of Innovations

Product development initiated through already existing technologies, and targeting established markets is under normal conditions described as continuous innovation. More so, continuous innovation is associated with products that have been altered by upgrading the underlying components (Pellissier, 2008). Further, Pellissier (2008) states that the continuous innovation is connected to improvements of existing components, and not changes to the products.

According to Garcia and Calantone (2002), discontinuous innovation is designation of radical change - innovation that drives on technology or market discontinuities, or both. Discontinuous innovation involves a new set of components that are merged through a new product architecture (Pellissier, 2008). Other research distinguish discontinuous innovation from continuous innovation by referring to fundamental differences, and a change of existing routines in the organization (Gopalakrishnan & Damanpour, 1997; Damanpour & Schneider, 2006; Garcia &

Calantone, 2002). At the same time, continuous innovation entails incremental change, and is described as a variation in existing practices.

The scope of innovations (continuous and discontinuous) and their total as well as relative extent of innovativeness has been a subject of debate. Robbins and O'Gorman (2014) highlight that academic researchers and corporate managers often focus their efforts on exploratory discontinuous innovation, leaving exploitative continuous innovation in the shadow. Further, existing empirical research suggest that continuous innovation might not even be perceived as innovation by managers (Robbins and O'Gorman, 2014).

2.1.1 The Generic Innovation Process

In the literature there is a common idea that innovation should be viewed as a process rather than an outcome (Boer & During, 2001). This translates into the fact that isolated ideas and inventions need to be further implemented through a process of commercialization to be regarded as innovations. According to Damanpour & Schneider (2006), a generic innovation process has a three stage progression. The three stages of the generic innovation process can enable organizations to successfully identify, organize, and analyze the key success factors which decisively affect and drive innovation.

The three stages of the innovation process are: 1) initiation, 2) adoption decision, and 3) implementation (Damanpour & Schneider, 2006). In the initiation stage the organization identifies an opportunity or a problem that needs to be solved. Tidd & Bessant (2009) explain that the initiation stage (*they call it search phase in their innovation process model) should be focused on detecting indicators of innovation potentials in the corporate environment. Furthermore, the investigation and evaluation of prospective solutions and proposal of the adoption are mentioned to be key organizational activities under the initiation stage (Rogers, 2003). Following initiation stage is the adoption decision stage which is directed towards the decision-making of either accepting or rejecting the innovation proposal (Wolfe, 1994). Damanpour & Schneider (2006) state that the adoption decision stage requires senior management's involvement for evaluation of the proposed concept and, where an approval from senior management would directly lead to resource allocation. Subsequently, the final step in the process is the implementation stage. Rogers (2003) explains that the implementation stage incorporates implementation activities, such as acquiring key resources, establishing procedures and policies, in addition to assessing an analyzing the utilization of the innovation.

2.1.2 The Fuzzy Front-End

This paper will mainly focus on the first two segments of the New Product Development process, and what role Big Data has within those segments. Initially, the first segment is the Fuzzy Front-end, the initial stage where ideas are generated, opportunities are identified, and business concepts are defined. Later, the paper aims at the more formal and structured development phase, and takes into account the sequential Stage-Gate model and emerged Lean Startup method.

What is the Fuzzy Front-end?

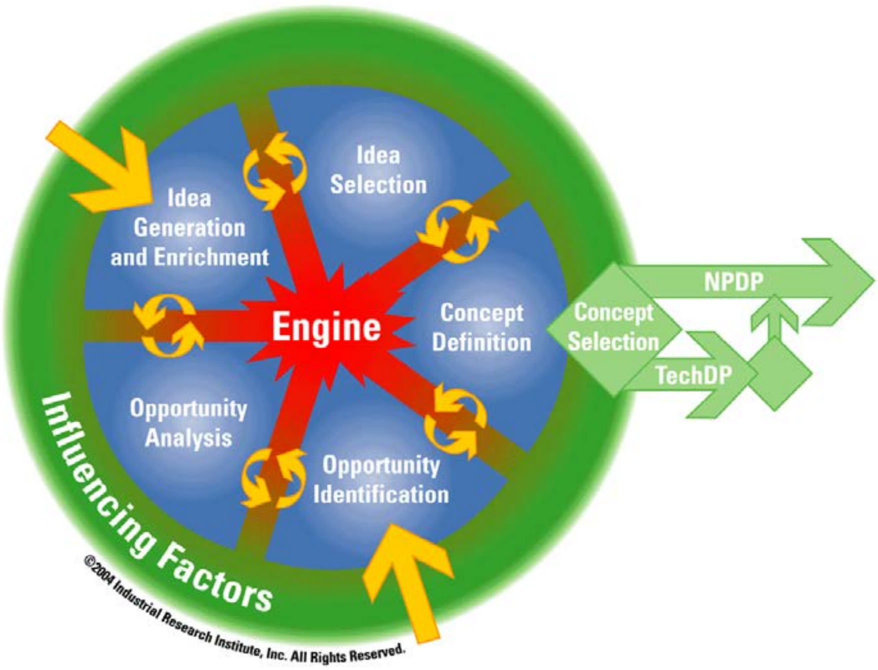
The Fuzzy Front-end (FFE) is the first and initial stage of New Product Development. The Fuzzy Front-end is the non-linear generator of opportunities, product ideas and concept which, if successful, are pushed into the formal and structured development process (Cooper, 1998; Moenaert et al., 1995; Stevens, 2014; Kim & Wilemon, 2002). Further, it incorporates the evaluation and decision-making about bringing the ideas and concepts into the formal development stage.

Being able to manage the pre-development stage, also known as “Fuzzy Front-end” of the New Product Development process, is considered to be one of the most important and challenging modules. Kim & Wilemon (2002) point out that effectively performed FFE activities can possess a direct influence on the success of a new product. Consequently, project as such are often seen as unsuccessful at the final part of the NPD process, or in other cases during the formal preceding stages of development. However, in many cases the source of failure comes from the pre-development phase. This statement is further strengthen by Cooper (1998) - who highlights the essentiality of the pre-development stage for product success. However, despite the importance of FFE, the early stages receive the least amount of resources in the NPD process (Cooper, 1998). Findings show that many companies identify that the weakness in their New Product Development process lies within FFE. Herstatt and Verworn (2001) point out that successful and high-performing products had significantly more (in some cases twofold) resources directed towards the fuzzy front end than unsuccessful products.

To generate high and successful performance in the FFE, organizations and managers need to understand the complex elements and the effects of the fuzzy front end (Kim & Wilemon, 2002). Nonetheless, the difficulty lies in the limited research and outlined structures of FFE and the pre-development stages in NPD. Kim & Wilemon (2002) argue that such limitation is grounded in the fact that the FFE stage is typically dynamic and holds no formal structures.

2.1.3 Elements of the Fuzzy Front End

The complex elements and activities of the FFE are visualized with the New Concept Development Model (NCD).



(Source: Industrial Research Institute, 2004)

Exhibit 1. An overview of the New Concept Development model (NCD)

The NCD model is composed of three key segments. The first segment is located in the center of the model which is the engine that drives the five components managed by the organization. The engine involves the internal environment of the organization, such as corporate governance, strategy and culture. Koen et al. (2001) argue that a proficient engine is essential to stimulate innovation. Referring to the fact that the innovation process needs to be aligned with the business strategy (strategic fit) and ensure continuous corporate leadership involvement.

The second segment is the outer layer which incorporates the influencing factors into the model. Belliveau et al. (2002) defines the influencing factors as organizational capabilities and resources, external macro environment (political, economic, social, technological, environmental and legal) and external micro environment (customers, competitors, distribution channels). It is improbable for organizations to control the influencing factors entering from the external environment. Yet, the engine (governance, strategy and culture) is able to alter their impact on the key activity elements (Koen et al., 2001). Further, Koen et al. (2001) explain that the influencing factors enable serendipitous discoveries of new ideas through their constant influence on individuals.

The third segment of the NCD model is the inner part of the model containing five activity elements (opportunity identification, opportunity analysis, idea generation, idea selection, and concept definition). The five activity elements of the FFE are not structured as a sequential process, but operate in a flexible order or combination through loop-backs, repetition and redirection (Belliveau et al., 2002).

The five activity elements of the FFE	
Opportunity Identification	Identify opportunities through a creative approach (mind mapping, brainstorming, lateral thinking), problem-solving approach (fishbone diagrams, process mapping), ad-hoc sessions, individual insights or senior management incentive.
Opportunity Analysis	Information gathering to support and transform ideas into specified business and technology opportunities. Through focus-groups, market insight, what-if scenarios, and experimental research. Further, exploration of strategic fit and extensive use of business intelligence and trend analysis.
Idea Generation	Evolution from opportunity to concrete idea. In this stage, the idea is being developed, modified and tailored through an iteration process. The process often involves collaborations with end-users, cross-functional teams, and other organizations.
Idea Selection	The activity of screening and selecting which ideas to undertake. Idea selection decisions are based on either individual choices or more formal product portfolio management methods. The restricted amount of information and insight causes inabilities for formal project selections. Implemented idea selections models are evaluating and assessing technical and market uncertainty, investment tolerance, organizational capabilities, and strategic fit rather than solely focusing on financial returns.
Concept Definition	In this stage, the building of a business case is completed supported by market research, customer insights, capital and funding requirements, industry analysis, and overall project uncertainties. The defined concept (business plan or project proposal) is the final output of FFE before resource allocation and NPD is initiated.

Table 1. Explanation of the five NCD elements (Source: Koen et al., 2001)

2.1.4 Managing the Fuzzy Front End

According to available literature on the FFE (Herstatt & Verworn, 2001; Kim & Wilemon, 2002; Koen et al., 2001; Alam, 2006), certain success factors go into effectively managing and organizing the pre-development stage. Such factors cannot be regarded as sovereign success factors, thus organizations have to consider the interdependency of them, and find favorable combinations. In the previous section, this paper introduced key activity elements of the FFE. Previous theory suggests specific settings and approaches of these elements to increase the possibility of product commercialization success.

An important factor for success in the FFE, is the presence of product champions - a senior individual in command of the internal development and external activities connected to the specific product. According to Heller (2000), products with product champions have a greater success rate, since champions promote the interpretation of product concepts, both in the internal- and external domain.

An organization's proficiency in concept development and concept definition has a positive correlation to product success (Cooper & Kleinschmidt, 2001). According to Bacon et al. (1994), proficiency in concept development and concept definition requires cross-functional evaluation and information sharing. Bacon et al. (1994) emphasize that results from analysis of information needs to be communicated across functions to enable the access of the organization's larger knowledge base. Furthermore, Bacon et al. (1994) identify the necessity of modifying the concept during the development. Such concept modifications can lead to ambiguity and interrupt the formal development, but necessary for adjusting the concept to suit technology- and market changes.

Having the right degree of formalization in the FFE and its activities is essential for success. Kim & Wilsmon (2002) discuss the absence of formalization in the FFE and how it translates in managerial challenges for organizing it. Other literature agree on that formalization is necessary for effectiveness, however, too much formalization and dependency on it can hurt the innovativeness in FFE. Hence, organizations need to consider what level of formalization is appropriate for their specific objectives and strategies in the FFE (Boeddrich, 2004; Khurana & Rosenthal, 1998; Gassman et al., 2006).

Bacon et al. (1994) addresses the importance of the NPD processes being aligned with general business strategies. Thus, indicate that product developments must take advantage of the dominant capabilities of their internal organizations. Khurana & Rosenthal (1998) further emphasize that product strategies need to be aligned with development processes to guide and make decision-making in the Fuzzy Front-end more efficient. Thus, a significant factor for success is the ability to establish alignment between business strategies, product strategies and decision-making, i.e. strategic fit.

Support and involvement from Senior Management in the FFE is considered to be a significant success factor. Kim & Wilemon (2002) argue that FFE performance will eventually decline with insufficient support and involvement from the senior management, since support and strategic alignment in FFE is needed to encourage successful concept development and NPD.

According to Koen et al. (2001) and Kuhrana & Rosenthal (1998), FFE activities are performed in small cross-functional project groups, for that reason Senior Management involvement is necessary for linking key element activities throughout functional boundaries.

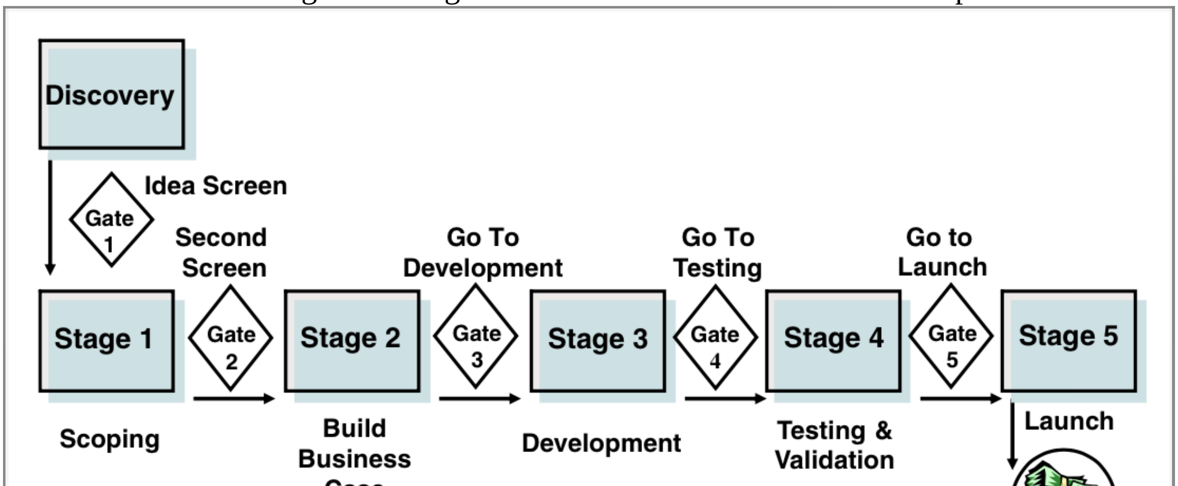
Cross-functional integration holds significant importance when structuring the FFE stage. Kim & Wilemon (2002) explain that organizations with a high level of cross-functional integration are able to support the New Product Development process more effectively, and adapt to change. In addition, cross-functional accessibility and sharing of information is essential for FFE. However, a multifunctional involvement is not considered a success factor by itself. Effective communication within the cross-functional setting is a key criteria for driving key element activities (Bacon et al., 1994).

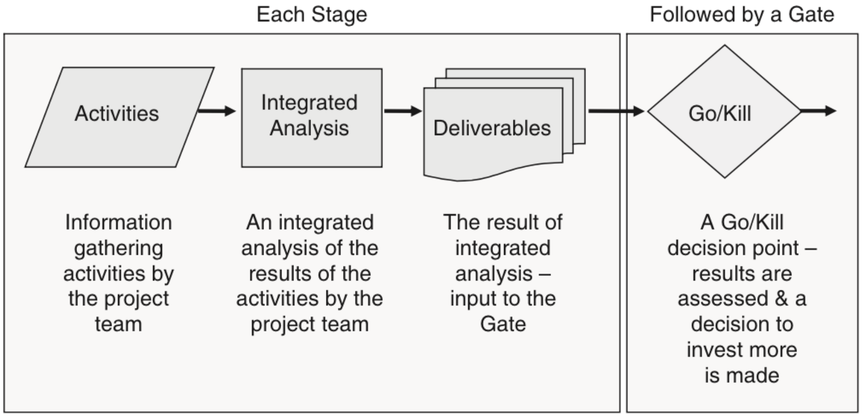
Kohn (2005) states that when it comes to idea identification and selection, organizations have to use exploratory rather than confirmatory methods for identifying and selecting new product projects. Kohn (2005) further argues that the purpose of the screening and selecting activities in the FFE is to explore new opportunities, rather than to reduce managerial- and technological uncertainty on New Product Development. Langerak et al. (2004) highlight the importance of customer insights and market orientation in the idea generation and idea selection activities. Hence, understanding of customer information and their needs is directly related to the effectiveness of generating new product ideas. Alam (2006) marks that allowing for influences from the external micro environment, i.e. customers, can aid selection and analysis activities, as well as reduce product development cycle times. However, Alam (2006) argues that activities should not completely rely on such influences, due to customers' inability to fully report their necessities.

2.1.5 Stage-Gate Model

After the initial stage of FFE the project enters a more formal sequential process of new product development. The Stage-Gate model is a value-creating process and risk model used to efficiently and profitably transform new ideas within an organization into successfully commercialized products (Cooper & Kleinschmidt, 2001; Cooper & Edgett, 2010). The Stage-Gate model involves the complex process of New Product Development (NPD) with a cross-functional view. A process is separated into different stages (where specific project activities are performed) and gates (where evaluation of completed activities and so called "go or kill" decisions are made, decisions concerning if the organization wants to pursue with the project or eliminate it). The complete Stage-Gate process is compiled by activities in the pre-development phase, the development phase, and the commercialization phase (Cooper, 2001).

Exhibit 2. An overview of a generic Stage-Gate model for New Product Development





The Stage-Gate model in Exhibit 2 is an overview of a five stage and five gate process, including discovery/generation phase and post-lunch review of the product. Cooper (2008) explains that each stage incorporates a set of required activities that are essential for the project to progress to the next gate. Each stage is cross-functional. In other words, activities of each stage are performed concurrently by professionals from different functions (marketing, R&D, finance, engineering etc.) within the corporation. The plan of each stage is also to gather information to reduce the risks and uncertainties of the projects, so that the risk can be effectively managed (Cooper, 2008).

Each stage is followed by a gate or so called go/kill decision point (Cooper, 2008). The gates in the Stage-Gate model have similar elements and structure. They consist of 1) Deliverables - that are delivered to the gate, in form of completed tasks of the previous stage. 2) Criteria - checklist of important requirements that need to be fulfilled. Criteria is mainly used to detect weak links and prioritize the projects. 3) Output - a decision with directions for the following stage (deliverables, timeline and required resources).

Exhibit 3. An illustration of Stage-Gate model's Stages and Gates (Source: Cooper, 2008)

2.1.6 From Development to Market-Launch

According to Koen et. al (2001), an approved concept in the Fuzzy Front End enters the structured development stage (as shown in the Exhibit 4 below, and in Exhibit 2 - it is represented by the third stage in the Stage-Gate model).

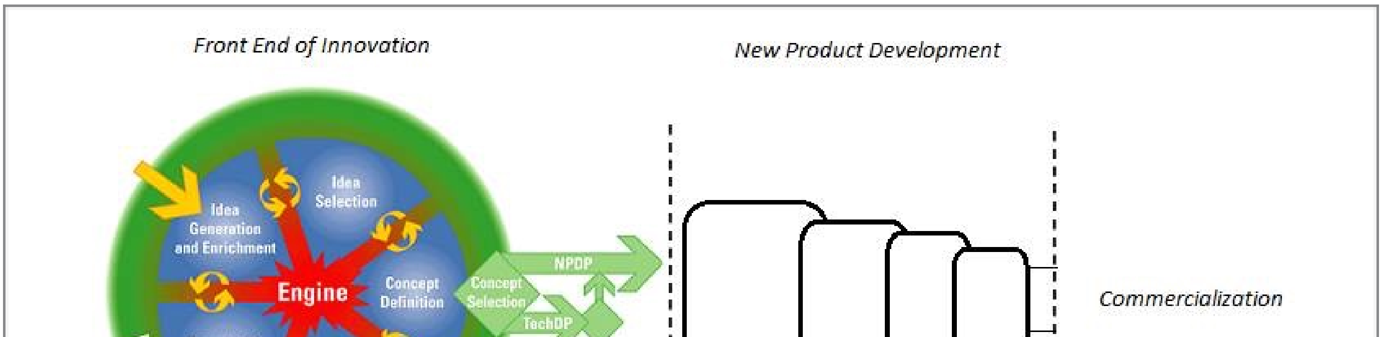


Exhibit 4. Three parts of the innovation process according to Koen (2001)

The development phase focuses on the technical tasks in parallel with manufacturing and marketing activities. Cooper (2001) explains that during the development stage market and production plans and programs are established. Further, the innovation project team reviews and adjusts the financial and legal analyses. The final outcome or deliverable from the development stage is a product prototype, which will go into the subsequent gate and potentially enter the Testing and Validation phase (Cooper, 2001).

During the Testing and Validation phase, the entire project and prototype is carefully examined in several aspects, such as the viability and desirability of the product. Activities during the pre-commercialization phase include; lab product trials to test quality and performance, simulation of product usage through real user trials, testing the production process, assessment of the launch/commercialization plan with market and consumer tests, and revision of the financial viability (Cooper, 2001; Cooper, 2008).

After the testing and validation phase the product if approved moves into the final stage - the commercialization which entails full production and market launch. The tested commercialization or marketing plan is formally implemented. Further, full production of the product is initiated and continuous operational- and financial risk actions are applied.

Benefits of the Stage-Gate model	***Source***
Comprehensive outline of duties and deliverables	Nasierowski (2008)
Control of the project development cycle	Nasierowski (2008)
Mitigates the risk of commercialized failure and increases the probability of successful launch	Nasierowski (2008)
Gatekeepers reduce project favoritism and discrimination	Cooper (2001)
The model is flexible and can be adapted into different organizational structures	Cooper (2001)
Reduction in lead times	Dodgson et al. (2008)
Effective product portfolio management	Dodgson et al. (2008)
Cost control at every gate	Lotz et al. (2009)

Table 2. Benefits of the Stage-Gate model

Implementing a Stage-Gate model in project development process, offers valuable advantages and raises several attributes. According to Nasierowski (2008), a Stage-Gate model because of its structure can offer a comprehensive outline of duties for human resources at each phase and at the same time it is easier to quantify the deliverables. Consequently, it serves towards a more flexible control of the project development cycle and acts as an incentive to significantly reduce lead times (Dodgson et al, 2008; Nasierowski, 2008). This flexibility is a driver for such models to be adapted into different organizational structures and successfully manage cross-functional projects (Cooper, 2001).

Focusing at each gate's positive outcome impacts, Lotz et al (2009) emphasizes the existence of cost control at every gate the importance it bears on controlling the budget of the project. Furthermore, gatekeepers have the role of evaluating the project while at the same time they avoid project favoritism and possible discrimination (Cooper, 2001). As a result, Nasierowski (2008) and Dodgson (2008), claim that the Stage-Gate model approach provides the necessary tools and incentives for an effective product portfolio management by mitigating the risk of commercialized failure, while at the same time it increases the probability of a successful product launch.

Limitations of the Stage-Gate model	***Source***
More suitable for incremental innovation	Sebell (2008)
Risk of "false negatives"	van den Bosch & Duysters (2014)
Difficult to kill projects in succeeding stages due to investment escalations	van den Bosch & Duysters (2014)
Friction between creativity and organizing	van den Bosch & Duysters (2014)
Requires significant resources	Nasierowski (2008)
Lack of top management involvement	Sebell (2008)

Table 3. Limitations of the Stage-Gate model

The Stage-Gate model is argued to be favoring only incremental innovation and product development (Sebell, 2008). Discontinuous innovation development is shown to be effective with more dynamic models than the Stage-Gate model (Bessant et al., 2005). Nagji and Tuff (2012) further explain that the downside of the Stage-Gate model is that the funnel process is unsuitable for radical innovations. It is difficult to predict the success of commercialization for discontinuous innovations and using the Stage-Gate model for that purpose can lead to "false negatives", where propitious innovations are prematurely eliminated. On the opposite, another limitation is connected to the model's structure, and reluctance to "kill" projects in succeeding phases due to substantial investment (capital expenditure) escalations (van den Bosch & Duysters, 2014).

Innovation development needs strong support from the top management, hence Stage-Gate model's cross-functional project teams are solely not sufficient for successfully developing and launching breakthrough ideas (Sebell, 2008). Nasierowski (2008), indicates that a positive outcome generated from the Stage-Gate system is dependent on significant commitment and therefore requires resources, which top management can see as problematic. Further, van den Bosch & Duysters (2014) argue that the closed system constraints corporate creativity, and intrapreneurship. In other words, there is a trade-off between strict process organization and internal creativity.

2.1.7 The Lean Startup Method

Empirical findings of this paper show that more and more organizations focus their efforts and adjust the NPD processes according to the Lean startup method. Accordingly, Big Data is an important factor of this method.

The Lean startup is an alternative method to traditional product and business development which was introduced by Ries (2011). Contrary to the linear process of traditional product development, the Lean Startup is an agile development with short and repeated cycles (Blank, 2013). Lean Startup method consists of 5 principles: Entrepreneurs, Management, Validated Learning, Innovation Accounting, and Build-Measure-Learn. The dominant objective of the Lean startup is to shorten product development cycles and develop products in accordance with customer needs. The method is related to the lean manufacturing approach by aiming to eliminate wasteful activities and increase value-adding actions in the product development process. Ries (2011) argues that the Lean startup method allows companies to succeed with less funding and iteratively building minimum viable products, with just enough set of features to realize the product to early customers and receive valuable input, also known as validated learning. The agility of the method is based on continuously receiving customer feedback throughout the product development process, thus being able to avoid committing resources to unwanted features that do not address customers' needs and reduce market risk (Ries, 2011).

Principles of the Lean Startup Method	
Entrepreneurs	Entrepreneurs are everywhere, for that reason the Lean Startup approach can be adopted by every company, regardless of size or industry
Management	Entrepreneurship requires effective management to create new products and services

Validated Learning	Learning by testing ideas through experiments and feedback
Innovation Accounting	Focus on measuring progress, milestones and prioritizing which requires a new design and approach
Build-Measure-Learn	Turn concepts into products, measure customer feedback and learn if to pivot or continue

Table 4. Principles of the Lean Startup Method (Source: Ries, 2011)

Adopting a customer development approach allows for testing and validating products and solutions through feedback loops, referred to as the Build-Measure-Learn feedback loop (Ries, 2011; Blank 2013). The Build-Measure-Learn (BML) feedback loop is considered to be the central part of the Lean startup method. According to Ries (2011), BML feedback loop is explaining the series of actions that are needed for an agile product development. The loop is a repeated process of building/tweaking a minimum viable product, measuring customer feedback, and learning from the received information to understand if the product concept is sufficient or if to pivot the concept.

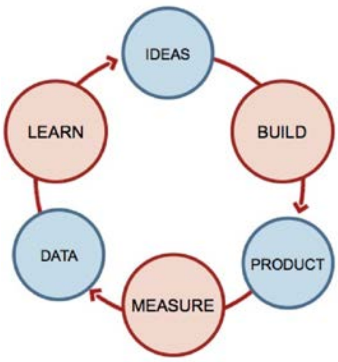


Exhibit 5. Build-Measure-Learn (BML) Feedback Loop (Source: Ries, 2011)

The Lean Startup Tools and Methods

Pivot is a method of evaluating and changing the strategic direction of an organization, after the realization that the progress made so far is insufficient or aiming towards the wrong objectives. (Ries, 2011). Further, Ries (2011) argues that the Lean startup is not offering a scientific procedure of performing a pivot process, but rather an alert to organizations regarding the persevere decision.

Ries (2011) explains that in order to proceed with the product development, designers can make use of certain tools. A/B testing is a process in which the developer is using two customized versions of the product in order to examine the market respond. Consequently, product developers are able to determine which version of the product is more desired by users and apply the improvements suggested by feedback loops.

Furthermore, Ries (2011) is using the terms actionable metrics and vanity metrics in order to explain where business decisions should be based on. Specifically, actionable metrics describes some of the parameters that are important for determining the focus of the development operations. The classification between actionable and vanity metrics is dependent from the core activities of each company and the urge to satisfy users' needs (Ries, 2011).

For software development, Ries (2011) suggests the continuous deployment process, in which the software code is regularly adapted and positioned into production in order to include improvements through the loop. The ability to gain and integrate users' feedback in the product faster than competitors, can result to a competitive advantage. Despite the fact that the continuous deployment is currently applied in software products, the author believes that it is possible to connect the method with hardware activities. As an illustration, Ries (2011) is describing the importance of software in mobile phones, possibly swift production changes and the flexibility of 3D printing tools.

2.2 Big Data

What is Big Data? McKinsey Global Institute (Manyika et al., 2011) explains in their report "Big Data: The next frontier for innovation, competition, and productivity" that Big Data is referring to large data sets whose size has increased to the point where current software tools are not able to capture, store, manage and analyze them. A similar definition is given by Dumbill (2013) - "*Big Data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn't fit the structures of your databases architectures. To gain value from this data, you must choose an alternative way to process it*" (Dumbill, 2013, p. 1). Accordingly, the term "Big Data" is defined by Dumbill (2013) as the inability to process data sets with mentioned attributes, such as the size and its incompatibility with existing information systems, thus there is a need to adopt to advanced information technologies.

In contrast, Feinleib (2014) argues that technical definitions which are expressed by majority of authors are valid but disregard the actual value of Big Data. Thus, the author defines Big Data as "*The ability to capture and analyze the data and gain actionable insights from that data at a much lower cost than was historically possible*" (Feinleib, 2014, p. 1). In that sense, Feinleib (2014) argues that Big Data should be interpreted by the extent of its impact rather than the technological capabilities of processing it.

Business or data analytics is not a modern trend. Data has been used effectively in the past and it is constantly progressing. However, that evolution is significantly more obvious with the Big Data movement. So what is the difference between Big Data and traditional analytics? Big Data as well as traditional analytics search for extracting value-adding information from data sources that can result in competitive advantage. However, Manyika et al. (2011) highlight three main differences: Volume, Velocity and Variety.

Morabito (2015) argues that Big Data is characterized by four dimensions: Volume, Velocity, Variety and Accessibility. Volume refers to the disproportionate large sizes of data and the smaller data storages held by businesses. Velocity is connected to the dynamic nature of Big Data and the swift "movement" which in many cases coincides with real-time. The dimension of Variety is describing the different types of data, which include a mixture from structured, semi-structured and unstructured data. The last dimension, Accessibility, is referring to the ability or inability of acquiring these packs of data. Further, many researchers describe the dimensions as the 4 V's and include Veracity as an extra dimension (Ohlhorst, 2013; Sathi, 2012). Veracity concerns the truthfulness and the access to a complete set of data, which could affect its reliability.

Since the establishment of the four dimensions (4 V's), there is a debate in academia regarding data under the label of Velocity (Morabito, 2015). Specifically, dynamic data that change constantly needs to be captured and processed in real-time and such data sets are recognized as Digital Data Streams (DDS). Due to that difference in the process timing they should be distinguished from Big Data. DDSs may be used for purposes that require a speed in action, such as frauds, medical surveillance, crimes, social media, while Big Data are to be analyzed after they are stored (Morabito, 2015).

2.2.1 Drivers of Big Data

Sathi (2012) argues that there are primarily three reasons behind the Big Data era: consumers, automation, and monetization. A more detailed explanation mentions sophisticated consumers, describing the generation of substantial information usage. The information usage enables insights about the market, as well as the interactions in the social media to exchange opinions and reviews. Such digital interactions are beyond the limits of pictures and text, but they also include videos, sound clips and several multimedia tools which in many cases involve expert opinions and quality rating. The term automation is referring to all the digitalized means, such as mobile phone services, emails, social networks, as well as the modernization of older ones, which are capturing and providing an enormous pool of data for analysis. As a generalized source of data, Sathi (2012) categorizes it in three groups: the first group encloses physical electronic products, which are characterized as "smart" since they have the ability to retrieve data about various usage parameters. The second category is the electronic touch points of mobile devices, computers and other electronic devices. The last group is including all the components that are offering services or act as routes for the movement of data.

Another important driver, which set the way for the Big Data era, is monetization - the ability and incentive to gather, process and trade data for the use of potential new business development. Packages of data can be sold to interested parties and their value can change or even create a new market creating collection value. Most valuable attributes lie within personal location data, cookies, user behavior and usage data (Sathi, 2012).

A fairly new phenomenon which enables further growth of Big Data is called "Internet of Things" (IoT). IoT is mainly describing machine intelligence, where networked technologies or smart devices (products and services that are used on daily basis) are interconnected and communicate through software and sensors (RFID). IoT enables fast and continuous exchange of real-time data for improving functionality, processes, and discovering new and improved products or services (Xia et al, 2012; Kopetz, 2011; Gubbi et al., 2013).

2.2.2 Benefits of Big Data

Literature indicate that Big Data can unlock plenty of new opportunities, and deliver operational and financial value (Ohlhorst, 2013; Morabito, 2015; Sathi, 2012). For that reason, companies are devoting their resources and efforts to gain greater results by leveraging Big Data. Davenport (2014) describes three main benefits of Big Data analytics: 1) Cost efficiency & effectiveness, 2) Enhanced decision-making, and 3) Exploring new opportunities.

By implementing Big Data technologies (such as free open-source software, inexpensive servers and cloud-based analytics) companies can support the already existing data processing tools, which will result in cost reductions. Sathi (2012) illustrates that an implementation of Big Data technologies would eventually lead to reduced latency, and require less administrators, hence supporting cutbacks in resources. Previous empirical research reveal that large organizations adopt Big Data technologies with the objective to strengthen their traditional technologies, not to replace them (Davenport, 2014). Manyika et al. (2011), suggest that Big Data significantly improves efficiency by allowing companies to raise productivity or enhance product quality by increasing its value. The analysis of production data can result to an optimized use of resources such as time, human resources and raw materials. Ohlhorst (2013), argues that data could also improve the stages before and after the production of the supply chain. Furthermore, Feinleib (2014), adds that combining production data with data from different functions, can provide the analysts with vital information on how to improve efficiency and effectiveness.

Another benefit of Big Data analytics is improved and enhanced decision-making. McAfee and Brynjolfsson (2012) argue that data-driven decisions seem to be more informed and effective. Organizations can leverage Big Data analytics to be more effective and faster in their decision-making, as well as acquire new capabilities to make evidence-based decisions. However, Ross et al. (2013) suggest that adjustments of current processes and corporate cultures have to be considered for a positive effect.

Lastly, Davenport (2014) argues that the most interesting use of Big Data is for new business development, in other words to improve and create new products and services across the value-chain. Data-intensive companies such as Google, eBay, Amazon and Facebook are continuously uncovering additional revenue and new value-streams through Big Data analytics. However, not only IT and data-intensive companies engage in harvesting Big Data for value. Strong focus on Big Data is also obvious across more traditional sectors. More so, a research conducted by Capgemini (2015) show that the majority (53%) of large established companies anticipate an increased competition in the future from start-ups enabled by Big Data. Insights that the large data sets bring can transform business models, enhance innovation capabilities and productivity, and enable companies to discover new markets by data-driven market learning (Gobble, 2013; O’Connor, 1998; Chen et al., 2012).

2.2.3 Challenges of Big Data

As presented in the previous section there are many opportunities linked to Big Data analytics and implementation of such into concurrent business activities of organizations. In order to harvest real value from extremely large data sets organizations have to address and overcome decisive obstacles, connected to both managerial and technological contexts.

Morabito (2015; 2014) addresses technological challenges of using and implementing Big Data to create actual business value. One of the main issues is the incompatible IT infrastructures and data architectures. IT systems and software should be able to store, analyze, and derive useful information from datasets, composed of structured, semi-structured and unstructured data. Further, Morabito (2014) argues that an enterprise-wide platform of sharing Big Data and its analytics within the organization and its sectoral system imposes challenges due to incompatible technologies. Manyika et al. (2011) additionally indicate that a key obstacle is the consistency of internal and external databases, implying that there is a challenge in integrating and standardizing data of contrasting formats to enable valuable information flows.

Accessing the required data could be a difficult task for a number of organizations and it is common that the acquisition of data stems from external sources, such as third parties. A SAS Institute report (SAS Institute, 2013) is describing a number of barriers in regards to the use of Big Data, after its acquisition. Particularly, it emphasizes on the time and speed that data should be acquired and processed to avoid being outdated and have their value diminished. Furthermore, it is crucial to ensure that datasets are complying with two important criteria: understanding and quality. The first refers to the ability to comprehend and separate useful as well as relevant data, which will form the information we seek, rather than including unconnected and misleading data. Data quality could affect the final outcome in a similar way by providing disorienting information which is not transformed into the desired value. Consequently, the decision-making process could be negatively affected.

An additional challenge of Big Data usage is the output information management. Simon (2013) claims that the outcome of Big Data analytics should be meaningful information. The challenge underlies in the ability to extract useful and targeted clusters of information out of the information pool, which require an urgent utilization.

Besides from the many technological challenges Big Data brings there are also managerial challenges that needs to be considered. Morabito (2015) argues that a major obstacle to overcome is management's lack of understanding of the potential value Big Data can bring to the business. Manyika et al. (2011) formulates in their report "*Organizational leaders need to understand that Big Data can unlock value - and how to use it to that effect*" (Manyika et al., 2011, p. 108). At the same time, data scientist need to recognize the business aspects to be able to bring valuable information to the business (Harris et al., 2013). In consequence, there is a lack of resources in the form of technical expertise combined with business understanding to position favorable Big Data structures and generate value. However, an already existing problem is the shortage of deep analytical and technical talent needed to leverage Big Data. According to a report conducted by Accenture, "*The U.S. is expected to create around 400,000 new data science jobs between 2010 and*

2015, but is likely to produce only about 140,000 qualified graduates to fill them" (Harris et al., 2013, p.4).

Another significant barrier according to Morabito (2015), is related to the funding of an analytics unit for Big Data analytics. The cost of establishing advanced tools in order to sustain a data processing department could be excessive. Therefore, IT and business executives should determine what the real needs for analytics are, settle on a budget agreement, and provide the funds for specifically these functions.

Firms are required to deal with several legal issues to be able to seize the full potential of Big Data. Extremely large amount of data is being transferred across both public and private networks, resulting in the establishment of various data policies. Data policies address mainly legal concerns of intellectual property, liability, security and privacy. Manyika et al. (2011) emphasize the issues related to privacy and usage of personal data such as medical- and financial records. Personal data is considered as valuable insight for increasing utility, however the specific category of data is viewed as sensitive and corporations need to comply with data policies. Kerr & Earle (2013), are examining Google's plan to enhance its search engine with predictive algorithms based on Big Data. The "intelligent" search, will display results to users even before they know they need it. They raise the concern that this use of Big Data promises efficiency and profits but they are skeptical that it might be used as a justification for shifting policies. As a result of Big Data technologies and corporate objectives, society will have to position itself in the trade-off between privacy and utility (Manyika et al., 2011; Kerr & Earle, 2013).

Morabito (2014) states that security is a barrier for the Big Data technology. The law requests that customer data needs to be protected and sets a limitation on which personal data can be acquired and used. In some particular cases, policies can be extensively strict, such as medical and financial records (Ohlhorst, 2013). The challenge for organizations handling Big Data has to do with liability, in other words the constant danger of security breach and the unauthorized use of data. There have been several cases of data hacking in the past and according to Morabito (2014), several organizations are reluctant to report online attacks, driven by the fear of public image damage.

Vare and Mattioli (2014) bring attention towards another challenge of Big Data, which is Intellectual Property (IP). In their article "Big Business, Big Government and Big Legal Questions", they explain the limitations on Big Data usage under the current IP legal system. Further, the authors reason that effectiveness of Big Data lies in the availability and transferability of data throughout organizations and systems, but there is an interference from privacy and IP legal frameworks. More so, understanding the legal rights of single strings of data and larger datasets are additional challenges which have to be addressed. Data is not patentable and copyrightable, however can be protectable trade secrets if they bring economic value (Vare & Mattioli, 2014).

2.3 Implementing Big Data

The purpose of all Big Data initiatives is to exploit valuable combinations of large structured and unstructured data, being both internal and external (Bloem et al.,2013). As an example, the data gathered from a telecom firm and financial services establishment could through a trans-sectorial connection enable a new type of product or service suitable for a totally different market. Integration of data pools are therefore required for significant potential, however this does not solely focus on trans-sectorial data pools as mentioned above, but also the integration of internal cross-functional data pools (Manyika et al., 2011).

To enable Big Data potential organizations need to coordinate and collaborate with other major Big Data actors and users. Accordingly, organizations will assist each other in the development of their data intelligence, together with IT partners, active suppliers and analytical-tool firms of large data sets (Bloem et al., 2013). Further, Bloem et al. (2013) explain that selecting the Big Data network has to reflect the strategic decisions and the long-term vision of engaging in Big Data activities. Additionally, organizational transformation is essential in accordance to the ability to process Big Data with regards to technological capacity, the potential of Big Data platform to comprehend and make use of incoming external and unstructured data. This combination of technological capacity and organizational change or transformation is a major and important factor of accessing Big Data potential (Bloem et al., 2013).

2.3.1 Data

Data is a broad term which incorporates various factors such as quality, reliability, availability, accessibility, as well as relevance, governance, and security, of unstructured, semi-structured, and structured data and its information (see table 5).

According to research (DiFilippo & Blase, 2014), Big Data has changed the decision-making process. Changes in decision-making are connected to intensified usage of analytical technologies and techniques. Further, companies are establishing supportive data insight functions within their organizations to inform decision-making. Thus, strategic, operational, and tactical decisions are relying in a greater extent on the increased data analysis and predictive analytics through real-time data.

Real-time data refers to streams of data that are delivered directly after data collection. In short, there exists no delay in the transition from raw data collection to the information provided from real-time data (Ozzimo, 2014). Aside from real-time data, companies further commit resources to extract value from current and historical data. Excessive emphasis on real-time data can lead result in complications and failure of data-driven decision-making. Wengler (2001) argues that availability of current and historical data is necessary to position real-time data into context of patterns and trends. Further, “right-time data” and actionable data needs to be provided to support decision-making (Ozzimo, 2014). Hence, relevant data needs to be gathered and integrated for specific requirements or scenarios.

Types of Data	
Structured Data	• Table and Records
Unstructured Data	• Human Language, Audio and Video
Semi Structured Data	• XML and Similar Standards
Event Data	• Messages (usually in Real Time)
Complex Data	• Hierarchical or Legacy Sources
Spartial Data	• Long/Lat Coordinates and GPS Output
Social Media Data	• Blogs, Tweets and Social Networks
Web Logs & Clickstreams	
Scientific Data	• Astronomy, Genomes and Physics
Machine-Generated Data	• Sensors, RFID and Devices
Metadata	• Data that describes the content of other data

Table 5. Types of Data and Data Sources

2.3.2 Staff

Staff refers to technology and analytic skills within the organization's human resources, presence and function of internal collaborative networks, as well as overall organizational structures and culture that serve to enable Big Data potential. According to Manyika et al. (2011), a major challenge for organizations is the lack of analytic and technology skills to make use of the Big Data. However, Manyika et al. (2011) acknowledge that the most significant barrier to overcome is to establish a data-driven culture and organizational structure.

According to Galbraith (2014), before organizations can entirely employ, accept and extract value from the Big Data platform, they are required to empower and embed the data personnel into decision processes. Galbraith (2014) argues that for organizations to be able to achieve this it will require a shift in power from current decision makers to data decision makers. In addition, there is a need to strengthen the data acceptance, data reliance and analytic capabilities within the organization to achieve a data-driven culture. A data-driven culture sees data and analytics as a central function, and by promoting a cross-functional distribution of data it allows for more fact-based decision-making (Parmar et al., 2014).

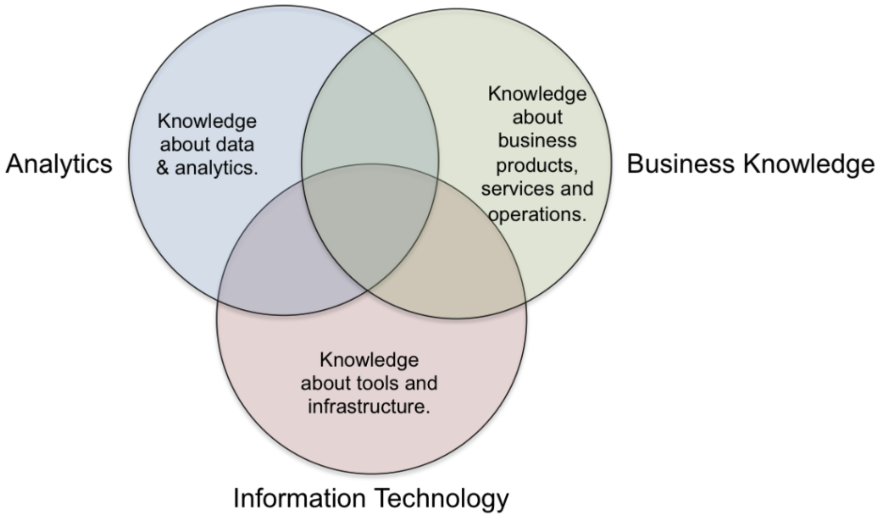


Exhibit 6. Knowledge fusion in a data-driven organization

Grossman & Siegel (2014), suggest that organizations have to unite information technology, analytics, and business knowledge (As seen in Exhibit 6). Further, the authors argue that organizations need to possess "a critical mass" of data scientists, with deep enough knowledge within the domain. Also, data scientists are required to have adequate business knowledge and understanding to be able to construct models that will generate business value. Lake & Drake (2014) and Harris et al. (2013), argue that the shortage of data scientists is due to the lack of required skill combinations for modern data objectives. For that reason it is suggested for organizations to form data analytic teams that include a mixture of data scientists, both generalists and specialists, with different skill combinations (Grossman & Siegel, 2014; Lake & Drake, 2014). Lake & Drake (2014) are comparing the current situation with the late 1980's and the realization of hybrid managers, a combination of a business manager with extended technical knowledge.

Another challenge concerned with human resources is presented by Barlow (2013). The author argues that organizations need to find solutions for integrating of new talent into their current structures or re-inventing structures to allow new talent to generate value.

2.3.3 Technologies and Techniques

Morabito (2015), explains that Big Data has a proportional evolution with the innovative technologies of the analytic tools required for it. The exponential size growth of data is an incentive for organizations to develop new innovative methods and technologies on how to manage data. Innovation is connected in a large extend with the ability to acquire and analyze data, in order to produce valuable information. However, different Big Data tasks have certain technologies and techniques that are optimised for each task (Morabito, 2015). Major technologies and tools for leveraging and managing Big Data are presented below and in the Appendix (see Appendix B).

Hadoop

Big Data, open-source platform and a frequent appliance of MapReduce (Sathi, 2013), with the ability to engage with several data sets (Ohlhorst, 2013), either by collecting sources of data for massive processing, or read data directly from the database and run intensive machine learning jobs. Hadoop's strength, underlie to the capacity of solving the most usual Big Data challenge, the efficient storing and accessing of large databases (Ohlhorst, 2013). Furthermore, according to Morabito (2015) Hadoop has an exceptional resilience to faults, by detecting and coping with failures at the application layer.

1010DATA Cloud Analytics

1010data has the ability to deal with semi-structured and machine data. It allows data integration, observing, reporting as well as several innovative analysis such as machine learning and statistical analysis. It is an effective technique to assist customers in scaling hardware and offers a centralized data management function (Morabito, 2015).

Actian Analytics

Morabito (2015), describes Actian Analytics as an open-source database incorporating a high-speed analytical database management system (DBMS), based on the synthesis of cloud services, data integration and analytical DBMS options. The SQL ETL (Extract, Transform, Load) technique can be used in order to gather data from multiple sources, modify them according to the business requirements, and move the outcome into a database (Feinleib, 2014).

Cloudera

Cloudera is a tool that aids Hadoop to establish performance, assess, control, security and reliability. It offers a large collection of software complements for uses like backup, recovery, monitor and audit. Furthermore, it distributes hardware for configurations for leading Hadoop providers, and it is used for a rapid analysis of large structured databases (Morabito, 2015).

2.3.4 Data Management Process

Data Collection

Data collection is the process of generating data with the use of digital means, e.g. by monitoring activities through sensors (OECD, 2014). Morabito (2015), explains that an important organizational barrier is in the process of collecting data. Data collected from different sources or different timing could be inconsistent due to a malfunctioning source. Furthermore, it is vital to generate the metadata in order to register the type and the method of measurement (Ohlhorst, 2013). Due to the growth of the volume of data, organizations are challenged to acquire new storage platforms and engage to intelligent data collection (European Commission, 2013). Lowering the cost of data collection is assisting in the establishment of better means for data processing (Feinleib, 2014). However, a major concern that organizations need to be aware of during data collection, is privacy (European Commission, 2013).

Data Extraction

The process of data collection data does not provide any value itself, unless it is actionable data - in a state to be analyzed (Morabito, 2015). Ohlhorst (2013) illustrates an example by comparing the results of extracting information out of a MRI (Magnetic Resonance Image) and a surveillance photo. Therefore, it is vital to use a process and extract the crucial information in a structured way and which is ready to be analyzed. Furthermore, another important aspect of collecting data is related with their validity and accuracy (Ohlhorst, 2013). However, there are several technical features which might not be included in some Big Data analytic units, such as error models and well-recognized constraints. Feinleib (2014) adds that data cleansing based on a cloud technology, could compose a great market opportunity for future development.

Data Integration & Analysis

Gathering and storing data into a database that no one will be able to access or combine, is unproductive. Data aggregation, requires an automatic process in order to manage a wide and efficient analysis (Morabito, 2015; Ohlhorst, 2013). It is crucial that data formation should be readable and resolvable by analytic systems in order to be queried, modelled and finally analyzed. Big Data, has an unreliable, dynamic and diverse nature, therefore for the analysis and mining it requires to be accessible and integrated (Ohlhorst, 2013).

"It is easy to see how Big Data can enable the next generation of interactive data analysis, which by using automation can deliver real-time answers" (Ohlhorst, 2013).

With this quote, the author suggests that computer intelligence will create queries to analyze and sort out Big Data, thus increasing data value.

Interpretation

Following the analysis part in the process, comes the interpretation of the results. This stage is expected to generate most of the value of the process, by providing insight, assisting decision making or information in the form of patterns (Feinleib, 2014). An effective interpretation requires the user to comprehend the results, assisted by the analytics platform tools. It is possible to automate a procedure and interpret the data with little or no human supervision. Ohlhorst (2013), explains that a very usual process connected with the interpretation of data, is the data provenance. Users add extra information about the nature of the data and the results, in order to enhance the process which can be used repeatedly for analysis, with different questions, factors or data sets.

Information Systems

As mentioned earlier in this paper, analyzing Big Data could prove successful factor to enhance the decision-making process (Davenport, 2014). The outcome of collecting, integrating, analyzing and interpreting the results, is the generation and access to information. Just as with data, firms have to establish an effective control over information in order to survive and compete (Laudon & Laudon, 2011). Furthermore, the authors explain the two perspectives of information systems' necessity, the technical and the business perspective. The technical perspective has to do with information management and assisting the process of input, process and output. The business perspective is dealing with the organizational challenges such as decision-making support, strategy and the firm's culture. It is essential for organizations to implement a strategy regarding the management of information, in alignment with their level of maturity, which includes size, operation width, and integration level (Laudon & Laudon, 2011).

System Configuration	
Centralized Systems	All activities of developing and operating take place at a domestic environment.
Duplicated Systems	They are domestically developed but operating in external autonomous locations.
Decentralized Systems	Each external unit designs and develops its own systems, regardless of location.
Networked Systems	Development and operations are integrated and organized across all units of the organization.

Table 6. System Configurations (Source: Laudon & Laudon, 2011)

SYSTEM CONFIGURATION	Strategy			
	Domestic Exporter	Multinational	Franchiser	Transnational
Centralized	X			
Duplicated			X	
Decentralized	x	X	x	
Networked		x		X

Exhibit 7. Strategy and System Configurations (Source: Laudon & Laudon, 2011)

2.3.5 Organizational Intent

Manyika et al. (2011), suggest that organizational intent should be to implement several new technologies and techniques, in accordance with the Big Data strategy, in order to extract value from their data. The level of technological adaption and Big Data platform selection will be affected by the data maturity of the organization, while technological innovation could have a key role in managing the constantly increasing volume of Big Data. Sathi (2012), adds that the approach that each organization will choose, is influenced by the size, limited capabilities and the urgency of each organization to evolve and transform by forming a competitive strategy.

As mentioned previously in this paper, adopting a Big Data analytics platform, can cause three significant operational impacts. It has the ability reduce the latency of moving and processing data, extend the storage capacity thus the potential to acquire more data, and consequently reduce the cost of acquiring and managing data (Sathi, 2012). Each organization has to decide the approach regarding the implementation of Big Data technologies. Sathi (2012), suggests three different strategies: the revolutionary-, the evolutionary-, and the hybrid approach.

Organizations choose the revolutionary approach in order to install a completely new Big Data analytics unit. This requires the organization to transfer the data to a new environment and fuse them with business processes. The revolutionary approach is adopted by data-driven companies, and while it is very effective, it is also costly and demands high technical capabilities (Sathi, 2012).

In the evolutionary approach, Big Data is inserted into the current Business Intelligence platform. This method is used by mature Business Intelligence companies, it has limitations regarding the process of data and the impact of their results, but it’s operating with a lower cost (Sathi, 2012).

The final approach, the hybrid approach, is described by Sathi (2012) as an addition of new technologies into the already existing Business Intelligence tools. The outcome of the merge, is a parallel operation of the analytics units, which enables a high-performance of both of them in their own environment. The hybrid proposal is regarded as an efficient way to gradually transform current technologies.

3. Methodology

3.1 Research Strategy

To wholly understand the managerial implications and concepts of Big Data in the field of innovation, an external analysis of case companies was chosen as an appropriate approach. The external analysis aims to explore how Big Data is managed to drive innovation in a corporate setting. Investigate what types of benefits and managerial challenges exist of using Big Data in innovation activities, as well as what key success factors lie within the organization to extract value from Big Data. Suitably, the main method of our research strategy is to conduct a qualitative research, emphasizing on an inductive approach with the purpose of theory generation (Bryman & Bell, 2011). The qualitative research will enable insight on strategies, capabilities, resources and processes used to successfully incorporate Big Data and analytics into the New Product Development process.

To obtain solid and reliable information, the qualitative data needs to be discovered and obtained from individuals that possess relevant expertise and experience, in other words, linked by setting to New Product Development and/or Big Data. Further, a qualitative research approach allows for adjustments of the method in the process of data collection. This type of flexibility is a useful feature, if corrections are needed for refining findings. More so, the critique of qualitative research is noticed and taken into consideration, such as the risk of interpretations of collected data being biased, and generalization (Bryman & Bell, 2011). The objective of this thesis is to give an understanding of the role of Big Data and key factors for implementation in different settings, and not to draw generalizations across organizations.

The research and its main purpose is regulatory, to study how companies utilize Big Data and apply it on the NPD process. Therefore, our research will commit to the functionalist paradigm, lead with a problem-solving orientation together with the aim that our empirical findings will deliver logic answers to or stated research questions. A functionalist paradigm, according to Burrell and Morgan (1979), is the dominant paradigm used for organizational studies. Since it is based on objectivism, its goal is to achieve a rational explanation of the usage of Big Data in the NPD process. The rational focus is expressed by pragmatic enhancement of positivism which will be tested through our hypothesis.

3.2 Research Design

The research will take a comparative design, since the aim of the thesis is to compare and gather data in context of a qualitative interviews and multiple case studies. The main reason for the comparative design is due to the novelty of the research topic, on how Big Data is managed and organized to assist the NPD process.

The comparative design incorporates the logic of comparison, which implies that we can understand the utilization, benefits and challenges of Big Data in the NPD process better when comparing multiple cases (Bryman & Bell, 2011). Also, conducting multiple case studies, an extensive understanding will be obtained to guide a comparative analysis. The comparative design will outline the data collection in a cross-sectional design format with the intention of better comparing different corporate settings, in addition better understand how Big Data is being used in different settings, exploring similarities and differences.

The framework will further allow comparison to comprehend specific organizational behavior and structures for implementing and organizing Big Data into the NPD. Further, creating categories of qualitative data will ensure comparability of findings. The multiple-case study improves theory generation. By comparing qualitative cases, the comparison will suggest relevant concepts for emerging theory (Bryman & Bell, 2011).

According to the literature review, there is a lack of research on Big Data connected to Innovation Management which in turn provides solid reasons to conduct a research of multiple cases. The decision to focus on multiple cases rather than a single in-depth is grounded in the aim to acquire a wider understanding of how Big Data is managed and what factors are essential for Big Data-driven innovations. Due to subject of time constraints, the aim will be to examine six case studies.

3.3 Research Method

In this section we will outline the main methods and approaches of our secondary data collection and primary data collection, data analysis, sampling and ensuring quality of the research.

3.3.1 Secondary Data Collection

For the purpose of this research, a conception of Big Data and a comprehension of its functions in an innovation context was required. The framework of this thesis had the grounded theory approach as basis, thus our objective was to develop the theoretical framework in tandem with the data collection. According to Bryman & Bell (2011), the inductive approach of the grounded theory forms strong connections and parallels between empirical evidence, theory and empirical analysis. The secondary data in this paper was primarily derived from existing journal articles, academic papers, corporate publications and books within the field of Big Data, NPD and Innovation Management. Relevant literature was explored and extracted from accessible electronic databases, mainly Scopus, Google Scholar, Wiley Online Library, Emerald Insight, and GUNDA. Keywords we used and combined to limit the search results were: *Big Data, Innovation, Analytics, New Product Development, Business Development, Organizational Design.* Apart from content accuracy and relevance, we sought to investigate appropriate literature after following criteria: citation frequency (estimate of research impact), literature currency (how current are the findings in the available literature), and audience relevance (is the direction and focus appropriate). In addition, in some cases we got referred to other useful and subject-relevant literature by our initial literature and their references (snowball method). Our secondary data was segmented and categorized under: Big Data & Analytics, Innovation & New Product Development, and Implementation of Big Data in New Product Development.

3.3.2 Primary Data Collection

The chosen method of primary data collection involved qualitative semi-structured interviews with professionals working to some extent with Big Data. As mentioned, the chosen research design to answer our research question is in the form of multiple case studies. Bryman & Bell (2011) argue that semi-structured interviewing is suitable for a multiple case study research since it gives sufficient amount of structure to ensure comparability.

An interview guide (see Appendix C) assisted the semi-structured interviews by providing coordination and a frame of reference with mainly open and general questions connected to specific topics: Big Data understanding, New Product Development, and Big Data in the New Product Development. The semi-structured interview approach provided us with great amount of flexibility and allowed the interviewees to freely express their experiences and opinions, which enabled greater and broader coverage of the investigated topic. Due to the iterative attributes of semi-structured interviews, we were able to vary the sequence of the questions, as well as adapt and form follow-up questions as we see fit or based on the direction of the replies (Bryman & Bell, 2011).

3.3.3 Sampling

During the research process we were hosted by University of Twente in Enschede, the Netherlands. The region is recognized for being the enterprising research region and home of many ground breaking high-tech companies. This enabled us to interact with interesting individuals within the university and region that work with Big Data both from technical and business perspectives, allowing us to deepen our understanding of the topic through more open and informal conversations. In addition, the university and our external supervisor assisted us in discovering and establishing contact with potential case-companies.

For the qualitative data collection we targeted companies that explored and used Big Data within various corporate functions, mainly product development. Also, companies that assisted other companies in realizing Big Data since they possessed great amount of insights by dealing with Big Data throughout different fields and sectors. We contacted 24 individuals (in 14 companies) that had a managerial role and involvement in product development, innovation strategy and/or Big Data analytics, i.e. innovation managers, senior consultants, data scientists and business developers. The contact information was found through the university network, but in the main using the combination of the social media portal LinkedIn and Google search. Our initial contact with the targeted individuals was via E-mail, where we presented ourselves together with a short presentation of our research and clarified what we sought to investigate. Of the 24 sent E-mails we received 10 positive replies where we were invited to schedule an interview. Later, we decided to not pursue with two interviews due their inapplicability towards our research objective. The outcome was 8 interviews within 6 companies, as shown in the table below. Two of them were conducted by using the video call application Skype.

Company	Role of interview participant	Channel	Location
Algoritmica	Co-founder, Strategy & Business Development	Video Call	Enschede*
Datafloq	Founder, Big Data Strategist	Face-to-Face	The Hague
Thales Group (Nederland)	Business Innovation Manager	Face-to-Face	Hengelo
TNO	Digital Technology Advisor	Face-to-Face	The Hague
TNO	Innovation Project Manager	Face-to-Face	The Hague
Philips	Senior Innovation Consultant	Video Call	Enschede*
Philips	Director	Face-to-Face	Eindhoven
Apollo Vredestein	Research Manger	Face-to-Face	Enschede

Table 7. Overview of conducted Qualitative Interviews

3.3.4 Data Analysis

The analysis of the secondary and primary data relied on the grounded theory. After each qualitative interview we pursued with open coding of our data collection to be able to analyze the data at the same time as collecting it. The data collection was examined, conceptualized, subcategorized into 9 topic groups, and placed into tables to obtain an overview of the key empirical findings. The same method was used for analyzing the secondary data collection and construct the theoretical framework. Coding of the data to enable comparison is a key activity within the grounded theory (Bryman & Bell, 2011).

Following the coding and categorizing of the primary and secondary data, we attempted to draw conclusions and compare findings from our case companies. A transcription of the interview recordings (2 interviews were not recorded due to sensitive information settings) was created. A revision of the transcripts was undertaken before connecting and analyzing the empirical findings with theory. This approach allowed us to be sure we did not fumble or misapprehend essential information and reassure we placed the information into right contexts. The key findings of theory and empiricism were then compared, analyzed and reflected. In other words, the output was analyzed to find if meaningful patterns were emerging through the code definition, code frequency and important combinations of such codes. More so, a cross-case analysis was performed to draw comparisons and conclusions of different corporate settings and levels of utilization. Our qualitative analysis followed the general strategy of “relying in theoretical propositions” which is preferred in our case according to Yin (2011).

3.3.5 Quality of the Research

3.3.5.1 Validity

According to Bryman & Bell (2011), a fundamental quality criterion is the validity of the research, referring to how well our research measures its main objective - what it was intended to measure. This research does not aim at generalization across industry, but rather explore and be aware of certain factors that are of great importance for using Big Data in the NPD process. In qualitative research there is a risk of receiving wrong or fabricated answers which we later base the empirical evidence on. This can be based in the fact that respondents are biased in their views of the potential of Big Data. The strategy of semi-structured qualitative interviews allowed us to address difficulties, such as misinterpretations or insufficient answers by restating or reformulating questions. A well formulated research question increased the internal validity.

The external validity is increased due to the data collection from individuals at senior and managerial positions which are in a direct or indirect way involved within NPD or Big Data. Further, Bryman & Bell (2011) argue that multiple case studies increase the external validity. In addition, transcribing the interviews as well as sharing the collected data for input with the interviewees, where we did not record the interview, increased the external validity.

3.3.5.2 Reliability

Bryman & Bell (2011) state that ensuring external reliability in qualitative research is difficult since it requires reproducibility, hence often requiring the same conditions and settings during the research study. The reliability of this research is increased through grounded theory and a well-connected theoretical framework to the empirical evidence. Furthermore, by doing tables and comparisons of the unique interviews we could observe distinctions. Establishing a high level of internal reliability is tricky when conducting a qualitative research since it is based on subjectivism and individual perception of what has been said. Since we were two observers during the interviews we could mitigate such difficulties and ensure internal reliability and inter-observer consistency (Bryman & Bell, 2011). Performing semi-structured interviews further enriched the reliability of the data collection and increased reliability of our research. We aimed to spot areas within the answers of the participants and add or modify some questions in order to achieve a further exploration of our topic.

4. Empirical Findings

4.1 Case Companies

In this section, the paper will present the case companies and the interview participants. The purpose is to give an overview of the organizations and individuals that took part in the research.

4.1.1 Algoritmica

Algoritmica is a Dutch IT- and data analytics company based in Utrecht, and operates in the sustainable energy sector. Their software product predicts asset failures using mechanical device data to decrease downtime and maintenance costs. The software company combines data and smart algorithms to allow organizations to increase asset performance, as well as make devices and infrastructures smarter.

The interview participant is the co-founder of Algoritmica. The interviewee has a data scientist background and is currently holding a managerial role by working within the fields of strategy, project management and product development.

4.1.2 Datafloq

Datafloq is a knowledge-driven Big Data platform with the objective to connect all different Big Data stakeholders within the global Big Data market. Datafloq provides Big Data strategy consulting and is developing a database of Big Data technology vendors to assist companies in the implementation of Big Data strategies and Big Data solutions. The company is based in The Hague, Netherlands.

The interview participant is the founder of Datafloq and a Big Data strategist. He has a marketing management background. Further, the interviewee has published a book on developing and implementing Big Data strategies.

4.1.3 Thales Group (Nederland)

Thales Group is a French multinational corporation based in Paris-France with offices in 26 countries, and operating in the fields of Aerospace, Defense, Transportation and Security. Their products vary from weapon systems to in-flight electronic systems, and their main revenue is generated by military sales. According to their annual results, the year 2014, Thales Group had revenues of 12.97 billion Euros and above 68,000 employees in 50 countries. Thales Group is the 10th biggest company as measured by the volume of arms sales.

Interviewee participant at Thales Group (Nederland) is the manager of Business Innovation. She is mainly responsible for innovation of the non-core business. Previously, she has been involved in strategy and production sales at Thales Group (Nederland).

4.1.4 TNO

TNO (*Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek*) is a nonprofit knowledge-based organization and research institute in the Netherlands that provides applied research solutions to corporations, businesses and public organizations. TNO has more than 3,000 employees and is represented in 7 countries (the Netherlands, Belgium, Canada, Japan, Qatar, Singapore and Aruba) with 29 offices which of 23 are located in the Netherlands. TNO focus on innovation and development within five areas: Industry, Health, Defense & Security, Urbanization, and Sustainable Energy. TNO has a major role in the growing international network of leading research institutes, corporations, universities, and other knowledge-based organizations.

The interviewee (1) is an advisor and researcher at the strategy department of TNO. His area of expertise is Digital Technologies such as Big Data and the Internet of Things. Further, he is involved in and focusing on data-driven innovation and the impacts of data-driven innovation.

The second interview participant at TNO works as a project manager and consultant, in the field of business processes supported with innovative ICTs. Further, Interviewee (2) at TNO coordinates the internal R&D on Big Data.

4.1.5 Philips

Philips (Royal Philips) is a diversified technology company focusing on areas of Healthcare, Consumer Lifestyle and Lighting. The company is headquartered in the Netherlands and has approximately 115,000 employees. Philips is a leader in acute- and cardiac care, home- and oral healthcare, energy efficient lighting solutions and applications.

Interviewee (1) is a Senior Innovation Consultant focusing on Innovation Management at Philips Innovation Services. Philips Innovation Services is the in-house consulting department where technical expertise is offered and utilized throughout Philip to consult and solve innovation challenges. In the past, he has been involved in management teams for consumer devices and production facilities.

Interviewee participant (2) at Philips is holding the position of director and department head of Smart Personal Spaces at Philips Research. The department is the hub for the creation of novel Digital propositions for personal and consumer use. The main focus lies within Smart Homes, Internet of Things and Big Data.

4.1.6 Apollo Vredestein

Apollo Vredestein is a part of Apollo Tyres - a multinational organization that produces high-quality tyres for cars, as well as industrial and agricultural machinery. Apollo Vredestein has large production facilities, together with the European regional office and the global R&D center located in Enschede in the Netherlands.

The interviewee at Apollo Vredestein is working as a Research Manager and has a professional background in physics and modelling. He has previously worked as a system architect at a leading IT-company. Currently, the interviewee is engaged in a project to implement Big Data in a larger extent into processes at Apollo Vredestein.

4.2 Findings

This section will outline the findings in structured tables under 9 categories: 1) Big Data definition, 2) Big Data objective, 3) Data Management, 4) Organizational setup, 5) New Product Development structure, 6) Big Data in New Product Development, 7) Big Data benefits, 8) Managerial Challenges of using Big Data in the New Product Development, and 9) Key Success Factors. Table 8 presents an overview of participating organizations and the interviewee positions, together with given labels which will be used in the presentation of the empirical findings. In addition, Appendix A will present the empirical findings in text.

Organization	Role of the interviewee	Given label
Algoritmica	Co-founder, Strategy & Business Development	**(A)**
Datafloq	Founder and Big Data Strategist	**(B)**
Thales Group Nederland	Business Innovation Manager	**(C)**
TNO	Digital Technology advisor & researcher	**(D1)**
TNO	Innovation Project Manager and consultant	**(D2)**
Philips	Senior Innovation Consultant at Philips Innovation Services	**(E1)**
Philips	Head of Smart Personal Spaces at Philips Research	**(E2)**
Apollo Vredestein	Research Manager	**(F)**

Table 8. Given labels to interviewees for the presentation of empirical findings

Big Data Definition

During the conducted qualitative interviews the participants were asked to define and share their understanding of the term Big Data. The following results within the empirical findings are based on their definitions, thoughts and comprehension of the term Big Data.

Interview participant	Answer
Algoritmica (A)	• *Observe function, performance and predict failure* • *Improve products and projects, ensure that parameters are set*
Datafloq (B)	• *Better understand of how products behave* • *Drive recommendations for new concepts*
Thales Group Nederland (C)	• *Achieve internal and external security goals* • *Identify a potential and a need for data security, and data that needs to be protected* • *Data transformed in security products and solutions*
TNO (D1)	• *Personalized products and precision* • *More prediction, contextualization and automation* • *Forge new value networks*
TNO (D2)	• *Uncover patterns that are hidden* • *Confirm the things you already know* • *Get insights and make predictions*
Philips (E1)	• *Extend products and services* • *Ensure consumer satisfaction* • *Improve product quality* • *Cost and efficiency in production* • *Discover new markets through data*
Philips (E2)	• *Accurate predictions, understand what users value and build something tailored to that* • *Use data for discovering and developing new innovations*
Apollo Vredestein (F)	• *Amount of data is that large that it allows for complete coverage of all settings that can happen and create useful combinations* • *Data that enables reliable predictions without knowing anything about the relation between input and output*

Big Data Objective

In this section the interview participants were asked about their organizations' objective of using Big Data and their expectations on the data-mining initiative.

Interview participant	Answer
Algoritmica (A)	• *Big Data is just data that has increased in size and volume* • *Data that needs scaled up technologies, advanced techniques and tools to manage* • *Data that is an important component of the product is Big Data*
Datafloq (B)	• *Combination of different internal and external data sources* • *Should be called mixed data*
Thales Group Nederland (C)	• *Big Data stands for enormous amounts of data that is opening up* • *Exponential and a more complex data world*
TNO (D1)	• *Data that allows you to see things in a larger scale* • *Do things with data that you could not do before*
TNO (D2)	• *Any data-related problem where the size of the data also becomes a part of the problem* • *Number of tools, methods and approaches that enables us to do things with data which we were not able to do before* • *Definition changes when technology advances*
Philips (E1)	• *Unstructured raw data from production* • *Really unstructured data from a large set of applications - less accurate*
Philips (E2)	• *Not just the amount but the difficulty in understanding of what is in the data* • *Complex with ambiguous, unstructured data from multiple sources*
Apollo Vredestein (F)	• *Currently no formal Big Data objective* • *A general objective of having good products and getting development costs down, where Big Data could be a useful tool to do so*

Data Management

In this section the empirical evidence describes the data management procedures such as collection, quality evaluation, utility and the assessment of what types of data are considered valuable.

Interview participant	Answer
Algoritmica (A)	• *All available data is useful if it is connected to the objective* • *Streams of data from sensors, performance indicators and maintenance logs* • *Robust techniques and scripts to ensure data quality* • *If you do not know what you are looking for, your Big Data project will fail*
Datafloq (B)	• *The usefulness of data depends on the analytic or business objective* • *Data quality is established using complete and accurate data including all variables and aspects* • *Poor quality of data or wrong algorithms lead to wrong insight*
Thales Group Nederland (C)	• *Own machine-generated data* • *Open and external data provided by customers for specific products* • *Social media data for trend analysis to capture moods, events and behavior* • *Quality purely based on an application basis* • *Prioritizing data from trusted sources*
TNO (D1)	• *The gathering and usefulness of data needs to be based on the objective and purpose of the project* • *Data quality is connected to the limitation of data and the anticipated results* • *Linked data - combine different datasets in open-format* • *Metadata can describe the quality of the data*
TNO (D2)	• *Collect external data or data you have not thought of, to correlate and combine with data you already have* • *Open data from collaborators through agreements* • *One school of thought is to get all the data, and second know what data you are looking for* • *Quality depends on application and requirements* • *Production data, data influencing the development process and additional undiscovered data*
Philips (E1)	• *No common approach on how to capture and collect data* • *Wide set of tools used to analyze data* • *Using production data, field data from products, and external data* • *Privacy laws causing data collection inefficiency* • *Need of good quality data and trained staff to process and understand it*
Philips (E2)	• *Capture and collect own product-generated data and external public data such as social media data* • *Brokering data occurs but selectively* • *First anonymize the data, then work with it, to comply with values, integrity and privacy* • *Determine the reliability of a data source by the accuracy of past predictions* • *Use sufficient correlation in production to find significant additions to quality and improve accuracy* • *Reject or delete data which violates privacy regulations or is unreliable*
Apollo Vredestein (F)	• *Make use of manufacturing data to link input parameter data to the output* • *Sensors can be a Big Data source to create relations and understand the product usage to enhance product development* • *Combining data with other product related industries can enable good things* • *No clear approach on ensuring data quality, assessment occurs often after the initial relations are made by removing false datasets for reliability*

	• *cleaner data from product development process than production*

Organizational Setup

In this section, the interviewees were asked about the organizational setup for managing the Big Data within an organization. Referring to structures and responsibilities of implementing Big Data into the system. Furthermore, identifying management's support and initiatives for integrating Big Data into the NPD process.

Interview participant	Answer
Algoritmica (A)	• *Big Data trend has increased the top management's awareness of data potential* • *Combining data is essential for value, but decentralized data is killing its applications* • *Accessible data and data analysis skills in all levels of the company for improved decision-making*
Datafloq (B)	• *Structured in such way that collecting and sharing of data is promoted* • *Departments should have access to all data to enable data-driven culture* • *Designation of a CDO (Chief Data Officer) - as a bridge between business and data technology* • *Top-down incentives*
Thales Group Nederland (C)	• *Big organizational changes, higher importance of offering consultancy services* • *Right people, right methods and right culture* • *Management boosting the data usage in cyber environments to proactively advise on threats*
TNO (D1)	• *Connect globalized platforms with decentralized areas* • *Accessible data via visualization techniques and sandbox environments to staff without data knowledge* • *People with domain expertise need to be able to ask questions that can be answered with data*
TNO (D2)	• *Flexible teams experimenting, testing and validating data* • *Teams with some autonomy, allowed to fail and break some process regulations* • *Data scientists and business individuals running the processes should change their thinking to common goals*
Philips (E1)	• *Market intelligence for analyzing and capturing trends and feelings* • *Internal and external digital connectivity* • *Departments have specific setups for using Big Data* • *Architectural function to track requirements* • *No large formal functions to promote Big Data*
Philips (E2)	• *Have a core entity managing the sales and market data* • *Installed department focusing on Big Data* • *For research and innovation it is necessary to have new approaches with data* • *Teams with application knowledge that are focusing on certain markets and collect specific data*

	• *Data scientists spread all over the organization, enabling competencies*
Apollo Vredestein (F)	• *Data initiatives are based on a project basis for different objectives* • *Established electronics groups (to establish automated analysis systems) that have experience with Big Data will support and implement Big Data analytics*

New Product Development Structure

Participants were asked to describe their organization's NPD process - how the organization structures and manages the different phases of the NPD process. This section is further concerned with exploring what tools and techniques the case companies use to evaluate and analyze their projects through the development phases.

Interview participant	Answer
Algoritmica (A)	• *Lean NPD process to avoid delays* • *Prioritize by using simple tools, revision control, Kaizen method and other project management tools* • *Shift the product as early as possible to the customer for validation & feedback*
Datafloq (B)	• *Simulation analysis is a key part to strengthen concepts and features* • *Big Data idea generation from a customer perspective, via market research and behavior analysis*
Thales Group Nederland (C)	• *Strict Stage-Gate model and a lean startup model, depending on the product or innovation* • *Disruptive systems in NPD for novelty* • *Security-related development process needs to be flexible* • *Boarders of the NPD are becoming vague with open innovation* • *FFE with a feeding stream from external and internal sources into the concept bank* • *FFE activities initiated from both bottom-up and top-down*
TNO (D1)	• *NPD needs to have an iterative approach* • *Data-driven NPD not very different from other NPD in structure* • *Core data technologies intersect with vertical products and demands* • *Bend horizontal technologies into vertical solutions* • *Start-up acquisition for data-driven innovation*
TNO (D2)	• *NPD life-cycles involve a funnelling approach, filter idea thought stage-gate process* • *Data-driven innovation and Big Data application development aren't that different from traditional NPD processes* • *Data-driven approach could be similar to the startup approach and the open innovation approach*
Philips (E1)	• *Usage of Stage-Gate model process* • *Structure of FFE varies based on type and size of product* • *Wide-spread of tools to evaluate and assess development, challenge is to fit with the problem* • *Small teams testing ideas in the market or consumer groups* • *Prove if solutions are good and feasible for mass production* • *Same mechanism for small as big products* • *Reference architecture is the key to start*
Philips (E2)	• *Initiative towards Lean-Startup method* • *First stage is liberal and encourages experiments* • *Every innovation has to go through a funnel, some fast and some slower* • *Research is mostly exploratory, next phase is to prove the principle and move further by proving the semi-product* • *Have product champions*
Apollo Vredestein (F)	• *Feedback loop which evaluates how the product is perceived and used based on user feedback* • *Stage-Gate process for creating a new product, a 2 year development cycle* • *Different function groups in the process, the end products are derivatives of a platform which is built up of concepts*

Big Data in New Product Development

The following section is examining if or how the case companies make use and implement Big Data into their NPD processes. In other words, what function Big Data has throughout the process and in which phases is Big Data used. Participants were also asked to rate their utilization level of using Big Data in the NPD process.

Interview participant	Answer
Algoritmica (A)	• *Product development dependent on data* • *Start from the End, identify the problem and evaluate the usefulness of solving it, then test with Big Data if valuable predictions can be made* • *More data equals better R&D*
Datafloq (B)	• *Big Data can serve every phase of the NPD process, however requiring different approaches* • *The utilization of Big Data in NPD depends on how data-driven the company is* • *Central data more used in the formal product development*
Thales Group Nederland (C)	• *Successful utilization and implementation into core products and markets* • *Not yet successful in utilizing Big Data for NPD directed towards new and emerging markets*
TNO (D1)	• *Product connectivity enables data feedback loops, which improves product design, product development and maintenance* • *Products without Big Data elements can still use Big Data to enable innovation processes* • *When the data is part of the product you get different dynamics of using Big Data in R&D and product development* • *More used for enhancing products and saving money than NPD*
TNO (D2)	• *For existing products it can improve efficiency and effectiveness* • *For NPD it is a process of discovery, what is the best way to implement it* • *R&D with Big Data from more sources, collectively discover innovations* • *Big Data used in very early stage to give insights* • *Built-in feedback mechanisms for late-stage product development* • *Big Data more valuable for technical applications rather than consumer applications* • *Big Data in NPD is great in theory, but no real examples in practice*
Philips (E1)	• *Identify a new need or new clients for existing or new markets* • *Big Data to track activity and other elements within the product - Smart City* • *Optimize products and enable growth* • *Privacy laws restrict certain usage of Big Data in the NPD* • *Big Data used in the NPD is not done on a large scale, still exploring the application of it*
Philips (E2)	• *Utilize data in research, and in FFE to generate new ideas that can be introduced to the market* • *Big Data just one element, the process will change a bit but in essence stay the same* • *Market data has been used for a long time, connected products that generate data and utilizing that data is new*
Apollo Vredestein (F)	• *Most used in the manufacturing or production phase* • *Big Data is seen as a tool for analysis over many projects and serve decision making in the product development process* • *Many concepts will be in electronic parts and Big Data can set up the structure, especially if the concept has a data generating function*

Big Data Benefits

This section is related to what benefits or potential the interview participants expect or experience from using Big Data in the NPD process.

Interview participant	Answer
Algoritmica (A)	• *Core feature to make predictions and improve products and development* • *Re-invent business models and exploit business opportunities*
Datafloq (B)	• *Reduce time-to-market* • *Adapt ideas, concepts and products on behavior to better fit with the demand*
Thales Group Nederland (C)	• *Big Data creates new markets so we can continuously develop products*
TNO (D1)	• *Enhanced market research and targeting of consumers* • *Create new products and improve the value chain with new insights* • *Cost efficiency, and targeting new partnerships* • *Strategic positioning in the new emerging value ecosystems*
TNO (D2)	• *Efficiency and effectiveness for existing products* • *Get insights that help you make predictions* • *Prescriptive analytics to improve decision-making* • *Validity through descriptive analytics*
Philips (E1)	• *Cost reduction and efficiency from Big Data* • *Exploring new opportunities* • *Use Big Data to make an impact and difference in society (smart city)*
Philips (E2)	• *Insights in digital form (user tests), understand what customers' feelings are etc.* • *Utilizing the potential of data to identify what you need to target*
Apollo Vredestein (F)	• *Big Data has a lot of potential for products or development that can be easily tested in different circumstances* • *Potential only in the bigger context, if you can generate enough data*

Managerial Challenges of using Big Data in the NPD process

In this section, the interviewees were asked to explain and describe what managerial challenges of using Big Data exist, as well as what managerial challenges their own organization is facing connected to Big Data.

Interview participant	Answer
Algoritmica (A)	• *Acquire necessary resources and skills for managing Big Data* • *Understand the business case, and not see data as sole potential of value* • *Change to a data-driven culture* • *Have enough data to avoid cold-starts in the NPD*
Datafloq (B)	• *Employees are reluctant to change* • *Departments' unwillingness to share disables data combination* • *Acquiring right expertise* • *Integrating data from silos and different departments into centralized systems*
Thales Group Nederland (C)	• *Big Data challenges are also connected with Internet of Things and Cyberspace* • *Challenges becoming more complex* • *To have a clear overview of decisions and impacts is more difficult* • *Need to take more risks to be upfront* • *Scarce human resources for the Big Data objective* • *Challenge of Big Data and usage of intangible assets (IPR)*
TNO (D1)	• *Scaling up new ideas is a challenge due to legacy systems* • *Change requires resources, agility and capabilities* • *Challenge of data ownership, interoperability issues, trust and privacy to share and connect data with external actors*
TNO (D2)	• *Find the right and qualified resources* • *Top management needs to understand and possess basic knowledge about data technology and its processes* • *A main challenge is the cultural clash between data experts and managers"*
Philips (E1)	• *Big Data is relatively a new concept for Philips* • *Other tools have to be used than we have been using previously* • *Management initiative to take action for legislation change* • *How to find efficient methods in drawing conclusions and decisions*
Philips (E2)	• *Getting the right data is the biggest challenge, often the data is not yet there and you have to generate it* • *Have an interactive attitude and culture* • *Acquiring the right people* • *Accessible and open information* • *Managers need to possess right soft skills*
Apollo Vredestein (F)	• *Have large amounts of data which is clean and rightly organized* • *Potential but at the same time and resource consuming to clean out the data bases to measure and analyze relevant data in a systematic way* • *Accessing data from the eco-system of suppliers and users* • *Staff need to work according to set procedures and the data needs to be clean* • *Need to have time and vision to implement the Big Data correctly into the process* • *Cultural issues need to be addressed*

Key Success Factors

The interview participants were asked about what key success factors or criteria exist for successful organizing for Big Data. Interviewees were asked to describe organizational and structural key success factors of using and implementing Big Data, both in general and in the NPD process.

Interview participant	Answer
Algoritmica (A)	• *Processes and project management, as well as experience using data in different projects* • *Building and understanding the business case* • *Having a data-driven culture* • *Establishing connections between data and business professionals* • *Data which is centralized and accessible for everyone*
Datafloq (B)	• *Having centralized data which can be easily accessed and combined* • *Acquiring external data sources* • *Training the staff to be more Big Data educated* • *Culture to involve data into every process in the organization*
Thales Group Nederland (C)	• *Multidisciplinary teams of specialists* • *Engineers that can oversee the entire scope* • *Understand that handling data has both direct and indirect influences* • *Oversee all stakeholders and their interests in the data*
TNO (D1)	• *Big Data objective and strategy* • *Access to the data* • *Speed and hyper-connectivity* • *Cooperating with different stakeholders with shared ambition and goal to be able to share data*
TNO (D2)	• *Acceptance and adoption by the users and people that are expected to work with the outcomes of the Big Data analysis for NPD* • *Right timing and right framing* • *Privacy laws that allow Big Data Innovation* • *Mental model for the acceptance of data innovation and data analysis*
Philips (E1)	• *Trained staff need to understand the data being offered and what to do with it* • *Make sure that the legislation is not prohibiting the good use of Big Data* • *Faster analytics and conclusions from Big Data*
Philips (E2)	• *The right data and the ability to do sufficient data testing* • *Sense of urgency in the organization* • *The right data science competencies* • *Be able to build your first proof points early* • *Culture - openness to experiment and explore new things*
Apollo Vredestein (F)	• *Prove a success story and have early adopters in the development team* • *Efficient analysis - have good data base access and reliable connections* • *Acquiring sufficient resources (human and technology)* • *Identify if the Big Data approach is suitable for the problem you want to solve*

5. Analysis

This chapter examines the empirics, and aims to draw comparisons between case companies and relevant theory. The analysis will make use of the coding table presented in table 9 to cross-analyze and class both primary and secondary data.

Categories	Core Category	Description of Core Category
Big Data Definition	**Definition**	A broad explanation of how Big Data is perceived.
Big Data Benefits Big Data Objectives	**Benefits**	Possible benefits from using Big Data in order to support and enhance the NPD.
Data Management Organizational Setup NPD Structure Big Data in the NPD	**Utilization of Big Data in the NPD Process**	In what way is Big Data collected, processed and assessed in order to be used in various models and phases in the NPD process.
Managerial Challenges	**Challenges**	Possible managerial challenges that arise and limit a successful implementation of Big Data in the NPD.
Key Success Factors	**Implementation**	Factors which greatly impact and determine the success or failure of Data-Driven Innovation.

Table 9. Coding Table with core categories

5.1 Big Data definition

This thesis examined the contemporary theory on Big Data and investigated its definition from multiple literature sources (Manyika et al., 2011; Dumbill, 2013; Feinleib, 2014). The results, based on the analyzed literature, appeared to have a high degree of homogeneity, as all literature sources referred to the technical aspects that distinguish data from Big Data. However, Feinleib (2014) is describing a different definition. Specifically, the Feinleib (2014) suggests that Big Data is about collecting, processing and acquiring insights from data, at much lower costs than it was possible in the past.

Considering the parameters of the term “Big Data” it is obvious that the term has an abstract meaning. An illustration of the definition paradox is observed by trying to quantify the word “Big”, but also by the constant adoption of new tools in parallel with the increase in data volume. Therefore, interview participants express a different perception of what Big Data is, and depending on their experience they have developed an alternative explanation regarding the dimensions or definitions of Big Data.

This study was conducted with contributors from diverse organizations, the majority of which utilize Big Data for commercial purposes. Therefore, the empirical findings are connected with elements based on a business perspective and primarily with the functions of production, marketing, and Research & Development. As mentioned earlier in this thesis, the goal is to examine the utilization of Big Data in the NPD process. Accordingly, the thesis will not go into depth with technical concepts related to Big Data. Participants repeatedly referred to the mere size of the data, and thus it is still the dominant criterion in defining Big Data and in accordance with the examined literature. Furthermore, the adoption of new tools and scaled technologies was consistent with the literature as a significant factor to distinct between data and Big Data. Also, the description from participants was associated with the dimensions explained in the theoretical framework, volume, velocity, variety, veracity and accessibility.

However, interview participants expressed they own coherence and perception elements of Big Data. Specifically, interviewees (E2) and (B) referred to Big Data as mixed data from multiple internal and external sources, open and private data. In addition, participants (E1) and (A) believe that another data characteristic that enables them to use the term “Big Data” in their organization, is the connection of data with the product or the production process. (D1), (D2) and (F) added that data that allows them to do things with it that they could not do before, falls under the label of Big Data. Finally, participant (D2) expressed the belief that the term Big Data should be used when data size becomes part of the problem of managing it.

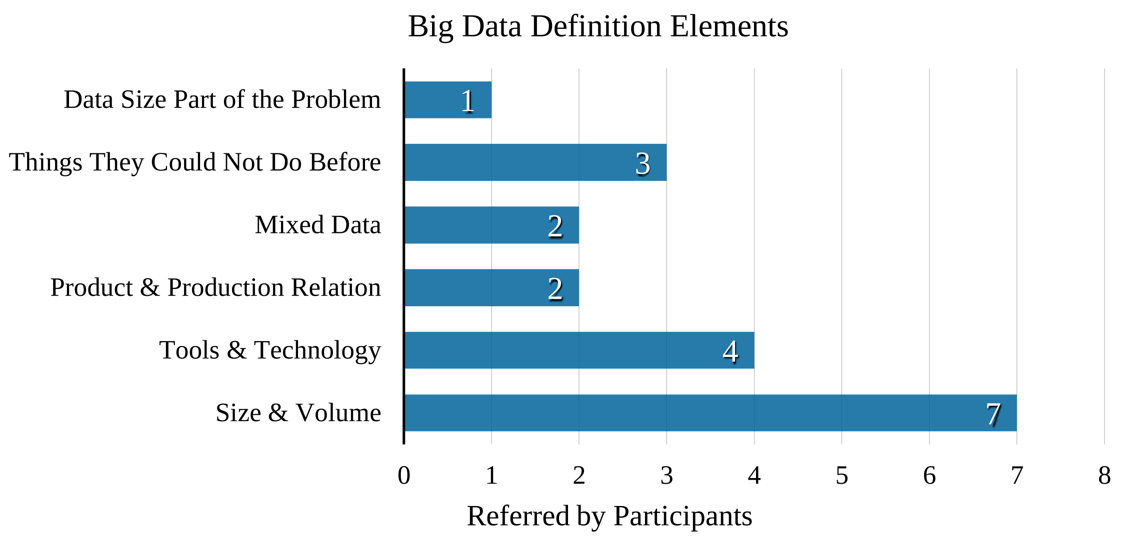


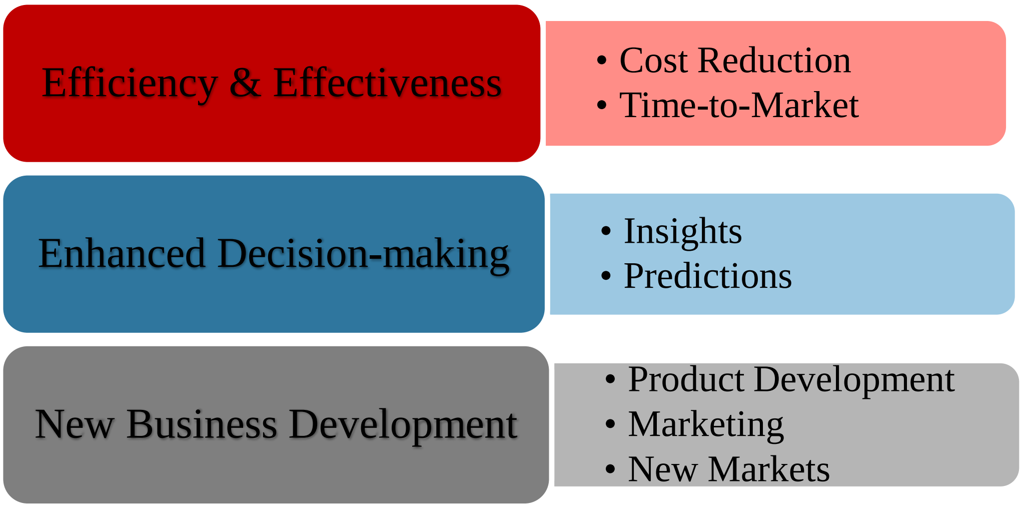
Exhibit 8. Characteristics defining Big Data within the empirical findings

Throughout the research, both the theoretical framework and the empirical findings indicated that it would be incorrect to define Big Data based on cardinal numbers. Thus, it is obvious that the size of data should be correlated with other parameters, such as the processing power and storing capacity of existing methods and technological capabilities. The analysis speed of these tools is one of the primary reasons for enabling Big Data potential in various business applications, including the NPD. Additionally, an important criterion of defining Big Data and distinction from traditional data, is the fact that Big Data is composed by a variety of data formats, which were not contained in the datasets in the past.

5.2 Big Data Benefits

In this part, the aim is to compare literature and empirical findings of benefits and objectives of using Big Data within the New Product Development. Since the existing literature on the subject was relatively scarce, the analysis will take into account the general potentials and benefits of Big Data and place them into context.

The empirical finding results appear to agree in a large degree with the theoretical framework. Specifically, the findings were placed into three categories: Cost Efficiency and Effectiveness, Enhanced Decision Making, and New Business Development.



Cost Efficiency and Effectiveness

Regarding the cost reduction and effectiveness of production, participants at TNO (D1, D2), Philips (E1) and Apollo (F) engaged in project or consultancy activities suggested that Big Data could be used to improve efficiency of existing products. Ohlhorst (2013) and Feinleib (2014), suggest that combining production data with data from different functions can strengthen efficiency and effectiveness in several stages within the NPD process, as well as optimize functions outside the NPD such as supply chain management. Also, participant from TNO (D1) explained that one way to accomplish cost reduction and increase productivity is through new partnerships. Finally, participant (B) from Datafloq added that Big Data implementation could improve effectiveness by reducing time-to-market.

Enhanced Decision-Making

McAfee & Brynjolfsson (2012) argue that data-driven decisions are more effective, since they are based on evidence provided by data. Most of the research contributors described the possibilities of how Big Data can improve the precision and validity of decision-making. These opportunities were grouped under two categories, Insights and Predictions.

Participant (B) at Datafloq, explained that utilizing Big Data can provide new or more precise insights by examining users' behavior and fitting it with demand. Interviewee (E2) from Philips added that insight can be in a digital form, such as use tests, in order to understand the customers and adjust the decisions about the products accordingly. Further, participant (D1) suggested that new insights can be used in the process of NPD for personalized products and precision, but also to enhance different stages of the whole value chain. His colleague at TNO, (D2) proposed that insights can be used to organize predictions in the market or in different sectors, and uncover hidden data patterns to use for commercial or social purposes. Additionally, (D2) said that he considers Big Data as a way of confirming expectations and decisions that organizations were already considering with the use of descriptive analytics.

The second label includes the responses from participants regarding predictions which can assist the process of decision-making by changing it into a data-based decision-making process. Particularly, interviewee (A) from Algoritmica argued that Big Data-based predictions can provide evidence for decisions, and improve a product which would be impossible to do so without using it. Contributor (E2) at Philips added that Big Data predictions could offer accuracy in sales, understanding what qualities user value and adjust products to it. Furthermore, individual (D1) emphasized that Big Data analytics can assist in the development of predictive, contextualized and automated solutions. His colleague (D2) at TNO explained that the prescriptive approach could enhance decision making but predicting and suggesting actions but it is currently in its infancy stage.

New Business Development

All case-study companies referred to the New Business Development process and how Big Data can complete the objectives set, when utilized in it. The information extracted from interviews was in accordance with the theoretical framework of this research. In order to provide a deeper explanation of the findings, we adopted three groups of New Business Development classification, Product Development, Marketing, and New Markets.

Algoritmica participant (A) suggested that utilizing Big Data in order to support the New Product Development process is a valuable tool to exploit business opportunities and re-invent business models. Additionally, he suggested that Big Data offer the possibility to observe and predict failures in product performance or its parameters, and avoid malfunctions or downtime. As mentioned in the theoretical framework, Chen et al. (2012) suggests that insights provided by data can transform business models and enhance the process of New Business Development. Datafloq contributor (B) believed that Big Data in the production process can lead to a greater understanding on how products or users behave, and enable accurate recommendations for existing or new products. Further, Thales interviewee (C) explained that Big Data compose a key element of some of Thales' products. By gathering data, they are able to determine product features and customize them accordingly. Davenport (2014) described the use of Big Data in the New Business Development process as the most interesting, since it enables products and services across the value-chain. Another participant, (E1) at Philips argued that Big Data contribute in the product development process by extending product features and life, as well as improvements in quality for customer satisfaction. Finally, he suggested that Big Data products could significantly impact societies, by offering products such as Philips' Smart-Cities concepts of connectivity and optimization.

Several interviewees identified the field of Marketing as another area of contribution of adopting a Big Data analytics unit in an organization. TNO participant (D1) suggests that Big Data utilization for marketing purposes can enhance market research, target consumers and the results can be translated into precision for creating more personalized products. Furthermore, contributor (E2) at Philips explained how Big Data can identify trends and customer demands, and uses their analysis to create and promote new product innovations. His colleague (E1) in Philips added that by using the above method, Philips can encourage customer satisfaction.

The last part of the New Business Development benefits that a number of participants identified in accordance to the theoretical framework, is the ability to acquire information about New Markets and Ecosystems. Interviewee (C) at Thales described that Thales is analyzing Big Data in order to monitor and determine potential need for new markets. After the examination of the results, the management team suggests if the organization should enter the market and create new innovative products. In Addition, contributor (E1) at Philips explained that his organization is exploiting Big Data in order to discover new markets and explore if these opportunities have the necessary investment potential. Finally, TNO interviewee (D1) argued that Big Data are a unique opportunity to strategically position inside emerging ecosystems and forge new value networks that could provide a competitive advantage.

5.3 Big Data and New Product Development

The Lean and Agile Approach

The empirical findings manifest that investigated case companies are focusing their efforts on being more lean and agile in the New Product Development process. Eric Ries (2011) introduced the new concept of product development - "the Lean Startup method", which is being utilized and implemented to some extent by both SMEs and large MNEs according to our findings. The Lean Startup is grounded on 5 principles with the objective to be more responsive to market needs by understanding necessities and eliminating wasteful activities of the NPD process. Findings at TNO show that data-driven NPD approach is similar to the Lean Startup method, as an iterative process of repeating activities and adjusting products based on the large amount of product relative data.

Algoritmica's NPD process is based on the principles coined by Ries (2011). Their agile product development approach is to iteratively develop a minimum viable product and shift the product as early as possible to the customer to enter the BML feedback loop. Interviewee (A) at Algoritmica explains that advanced NPD models will increase their product development cycles and set them back. Therefore, agility in the form of customer validation and other simpler tools are prioritized. Within their agile NPD process, Big Data is considered to be an essential component and enabler of validated learning and overall product development. Accessing and using Big Data in an agile development process enables them to iteratively modify and improve the product throughout the process. Further, as more data is collected and accessed the company are able to perform R&D on a larger scale to tweak, evaluate and develop with greater precision. The respondent (F) at Apollo describe elements of the lean approach in their NPD where user-feedback loops are used to improve products.

Philips, as a large and multinational enterprise, also recognizes the benefits of introducing the Lean Startup method for greater agility in the NPD process and a continuous customer validation and feedback to better fit the demand. However, according to our findings the Lean Startup method within Philips is only used in applicable situations and for suitable products. Thus, substantial and complex products (e.g. new X-ray machines) are developed through strict stage-gate processes. The findings show the same methods being adopted by Thales. Thales structures its NPD process depending on the type of product. Findings show that security-related product development needs to be flexible and adjustments have to be made throughout the entire development process. More so, Thales' market-related NPD takes basis in understanding and listening to the market, hence implementing an agile and lean process with high levels of customer involvement through feedback loops.

Data Collection

Research participants were asked to describe how do they collect and evaluate the data gathered. Morabito (2015) argues that collecting data is currently a limitation for several data-driven companies. Participant (A) from Algoritmica, identified the primary sources of data collection for his company as operations data, and specifically logs and sensors data. Interviewee (F) at Apollo emphasized the important Big Data source from sensors that can create relations and provide a better understanding of the product usage and in that way improve the development. Interviewee (C), described the data sources at Thales as own machine-generated data, open and external data provided by customers for certain product use and social media data. Contributor (D2) from TNO suggested a different approach regarding data collection. Specifically, he (D2) added that organizations should collect data they have not thought of in the past, in order to combine and correlate with data they already possess. Interviewee (E2) from Philips mentioned product data, field data and external data as the primary sources of collecting, but he argued that it is insufficient due to law restrictions. His colleague at Philips (E2) described the same sources with the addition that selective data brokering can occur selectively.

Data Quality

Furthermore, the accuracy and validity of the captured data is another challenging aspect that organizations need to address (Ohlhorst, 2013). Therefore, each organization should select and implement the effective tools, suitable for their purpose. Interviewee (A) from Algoritmica explained that the challenge of data quality is related to missing values and data has to be sorted with the usage of robust techniques and scripts. Participant (B) from Datafloq, suggested that data quality can be verified by complete and accurate data which includes values and variables relevant to the purpose of collecting them. Contributors (D1), (D2) from TNO and (C) from Thales, argued that the quality of data is connected with the anticipated results and based purely on application basis. As an illustration, participant (D1) described a situation where data quality is sufficient but not for the purpose which was acquired by the organization. Finally, interviewee (E2) at Philips connected the reliability with the accuracy of the past predictions, therefore he suggested as a method of accuracy improvement to correlate data in production in order to further connect it with quality.

Data Usefulness

Morabito (2015) suggests that in order for a process of data gathering to produce value, data is required to be actionable. Thus, each organization should seek data in accordance with the needs of its operations and goals. Algoritmica participant (A) believes that all available data connected any way with operations and products are useful, while any other kind of data is excluded. Interviewees (B) at Datafloq and (D1) at TNO, explained that the usefulness of data is determined by the analytic or business objective and if data is suitable for the right questions to be answered. (D2) at TNO suggested that he considers three types of data useful, production data, data influencing the development process anyhow, and undiscovered data that no one thought of but with value or an effect to the operations. Furthermore, he (D2) argued that they are two schools of thought regarding data usefulness, the first suggesting that all data are considered useful, and the second which suggests that organizations should acquire only the directly valuable for them data. Finally, participant (E2) from Philips explained that they consider all data useful as long as they are reliable and do not violate privacy regulations.

New Concept Development

As already described in this thesis, the New Concept Development model is composed by three segments (Koen et al., 2001). The first segment is describing the engine of the model, which includes the internal environment of the organization, and specifically the strategy and the culture. In accordance to the theory, participants (D1) at TNO and (C) from Thales suggested the importance of setting a data-driven strategy in order to unlock data opportunities. Thus, collaborating with stakeholders with shared objectives and interests offers an excellent prospect of expanding data-operations, rather than focusing just on data management. Furthermore, contributors (A) from Algoritmica, (B) from Datafloq, (C) at Thales, explained that a data-driven culture is the key for enhancing the decision-making process, and in several cases has a greater impact than factors such as the Big Data technology itself. Interviewees (D2) from TNO and (E2) from Philips, argued that data-driven culture could promote the adoption of Big Data in the New Product Development process by encouraging experimentation with data and encouraging a mental model for data innovation.

The second is referring to the influencing factors of the external micro and macro environment such as organizational capabilities and resources, as well as social, technological and legal elements (Belliveau et al., 2002). Participant (D2) at TNO and (E1) at Philips, suggested that social actors should position themselves in the changing reality of Big Data in order to promote the good use of analytics and data innovation. Therefore, the challenge underlies in encouraging the adoption of modern laws with achieves both a successful data commercialization and respect the privacy of sensitive user-data.

Additionally, all participants referred to the organizational capabilities for Big Data implementation in the NPD process. Bloem et al. (2013) emphasized the significance of organizational transformation in order to access Big Data potential. A crucial decision for the management team is to determine the information system strategy, according to the organization's level of maturity (Laudon & Laudon, 2011). Contributors (A) from Algoritmica, (D1) from TNO and (E1) from Philips, explained that accessibility and connectivity to all departments is able to increase effectiveness in collaboration and decision-making. Therefore, participants argued that centralized platforms can possibly prevent data-application elimination or critical delays, caused by decentralization of information systems, and no one supposed a decentralized model. In addition, interviewees (D2) from TNO, (E1) and (E2) from Philips, argued that designing autonomous departments and teams provide the necessary flexibility to support data experimentation and innovation. Contributors (D1) and (D2) from TNO, as well as (E2) from Philips, suggested the further involvement of data scientists within the organization in order to unlock competencies and enhance data processes. Finally, interviewee (C) at Thales added that apart from the people, organizations should make structural changes and implement the right methods and culture therefore organization consulting could be the key in enabling data capabilities.

The final segment of the New Concept Development model contains the elements within the flexible FFE loop, which includes opportunity identification and analysis, idea generation and selection, as well as concept definition (Koen et al., 2001; Belliveau et al., 2002). Thales participant (C) described the feeding stream towards the FFE as internal and external, bottom-up and top-down. This process is encouraging the flexibility in which the NCD model was based on. Furthermore, contributors (B) from Datafloq, (D2) from TNO, (E1) and (E2) from Philips, proposed that utilizing data in the FFE process, offer the ability to identify new needs and generate new ideas through market research and behavior analysis. Interviewee (E2) at Philips, argued that every data-innovation has to go through a funnel, in some cases it happens in the pre-development stages, consequently in the idea selection stage.

Applying Big Data into the NPD

The role of Big Data in the NPD process is directly affected by how data-driven the organization is overall. As seen at Algoritmica, their product and solution, as well as product development is highly dependent on the collection and processing of large data sets.

According to the theory and empirical findings, variants of the Stage-Gate model are the most utilized and preferred models within the NPD process. Both Philips and Thales make use of strict Stage-Gate models for advanced and complex product development. This approach is selected due to the many benefits the Stage-Gate model holds, i.e. strong control of the product development, mitigating the risk of failure, and other benefits which have been presented in Table 2 of this thesis. Datafloq (B) argued that Big Data can add value and be applied to every phase/stage of the NPD process. However, each case or purpose of the Big Data implementation requires a different approach and strategy. Interviewee (D1) at TNO emphasized that large data is more used for enhancing products and reducing cost than providing greater dynamics to the NPD process. Central data collected from the company’s own production lines or products’ inter-connectivity is used in the formal development phase.

During our research, we noticed that many companies are focusing towards more flexible NPD models or variations, such as the Lean-startup model. Such models can take advantages of Big Data implementation benefits in a higher degree, since their non-linear structure allows feedback loops and data-gained insights to be implemented as fast as possible. Participant (A) at Algoritmica argued that the initial goal of a new software product development is to receive customer insights by launching a minimum viable product, a process which aligns with the theoretical framework

(Ries, 2011). Interviewee (E2) at Philips, explained how the experimentation with data can enhance the NPD process in both Stage gate and Lean startup models.

We believe that Big Data have the ability to be used and transform all stages of the NPD process. However, we would suggest organizations to adopt flexible models or stages (NCD) in order to benefit even more. Linear models do not easily offer the ability to go backwards and redesign or implement changes. The more data-dependant a product is, the more it requires cyclical models to constantly improve. Feedback data is a continuous source of improvement, insights, and enhanced decision-making. To illustrate with an example, we can mention most software products that need to be updated constantly.

5.4 Managerial Challenges

The empirical findings acknowledged several significant challenges that are essential to addressed when using or implementing Big Data in NPD. The obstacles that need to be overcome are presented and divided into two categories: technical and managerial. The technical challenges presented by theory and empirics are more concerned with technical ability and functionality, which lands outside the scope of this research. However, to wholly understand what hurdles organizations face, we cannot disregard crucial technical challenges that possess a direct impact on the utilization of Big Data in NPD.

Technical	Managerial
• Incompatible IT	• Understanding Data Value
• Access to Data	• Corporate Culture Change
• Meaningful Output	• Human & Financial Resources
	• Legal & Security

Anchoring change in the corporate culture

A major managerial challenge of using Big Data in the NPD process recognized by most of the case companies was linked to anchoring change in the corporate culture. Algoritmica contributor (A) and (F) at Apollo, stressed that a significant obstacles faced by many that are trying to implement Big Data into their processes is that the corporate culture is not data-driven. Further, interviewee (A) at Algoritmica emphasizes that cultural change is often being neglected by domain experts that are too skeptical about what value Big Data can bring, as well as changing the already established processes. In the literature, Manyika et al. (2011) argue that the most crucial barrier to overcome is the establishment of a data-driven culture and structure. Datafloq presented a concurring opinion and reasoned that the challenge lies in employees being reluctant to change which creates significant hurdles since Big Data implementation is a prolonged and continuous process. Another managerial challenge mentioned by Datafloq, which further incorporates the corporate culture, is the challenge of internal competition and unwillingness to share data, information or knowledge across departments, thus, disabling a successful Big Data implementation. The findings were in line with the reasoning of Parmar et al. (2014), placing data and analytics as the central function for achieving a data-driven culture, hence the need of having a cross-functional distribution of data.

Interviewee (D2) at TNO presented another cultural aspect as a managerial challenge. He argued that the main challenge of implementing Big Data into the NPD process is the cultural clash between different types of key staff necessary for such implementation. The cultural clash is the inability of data experts and business managers to understand the common objectives. Galbraith (2014) emphasized the need to strengthen the data acceptance, data reliance and analytic capabilities within the organizations to achieve a data-driven culture. At Philips it was vital to establish an interactive culture and attitude so the cultural clash can be minimized by the staff being knowledgeable about the applications to ensure good interaction with data scientist and engineers. However, interviewee (E2) points out the importance of addressing the challenge of having or acquiring individuals that possess the right soft skills to be able to interact across their competence fields.

Incompatible Technologies

Theory in the area of Big Data mainly presented technological challenges of using and implementing Big Data. Morabito (2015) stated that an essential obstacle to overcome is to solve for incompatible technologies, IT infrastructures and Data architectures. He further argued that incompatible technologies offset the initiatives of sharing data across internal and external platforms. Datafloq referred to the challenge of integrating data silos into centralized systems. Accordingly, Manyika et al. (2011) pointed out the technological challenge of integrating and standardizing data of distinct formats. So, ensuring format consistency of internal and external databases is a challenge that needs to be addressed. According to the empirical findings, the implementation of technologies has to be compatible with the specific task, since some technologies are more suitable for certain objectives and goals. Interviewee (D2) at TNO expressed that the organization needs to adapt the necessary technologies based on the task and fit the implementation accordingly. However, this technological adaptation is subjected to the data maturity of the organization (Manyika et al., 2011). More so, TNO emphasized that legacy systems (outdated technologies) form obstacles and organizations are required to possess sufficient resources, capabilities and agility to successfully deal with such technological challenges.

Human and Financial Resources

In order to establish an analytics unit which will support Big Data processing, interested parties should be able to ensure that they hold the resources for the implementation. All the participant cases-companies referred to the necessity of securing the required resources. Four of them, participants (A) at Algoritmica, (B) at Datafloq, (C) at Thales, (D2) at TNO, and (F) at Apollo argued that employees should be more involved with data procedures. Further, literature emphasized the importance of resource acquisition and management, in order to successfully complete such a project (Morabito, 2015; Harris et al., 2013; Grossman & Siegel, 2014; Lake & Drake, 2014).

While examining the literature and during our empirical findings, we realized that the hardest resources to acquire are human resources. Therefore, we propose organizations to train their employees or alternatively, create multidisciplinary teams composed with individuals who possess the necessary skills to manage Big Data as well as generalists. Regarding the financial resources, we suggest interested parties to determine the level of engagement of their organization and invest accordingly. By doing so, budget teams will control the cost of the analytics project, and they will be able to compare the initial investment with the value of the information gained at later stages, before deciding further investments for Big Data.

Access to Data

Regarding access to data, we are referring to two different categories. The first category is describing the organization's ability to acquire data from internal and external sources. OECD (2014), and Morabito (2015), highlight the importance of being able to collect data. Interviewee (B) at Datafloq and (E) at Apollo highlighted the importance of establishing external sources for data, while participants (D1) at TNO and (F) at Apollo believe that stakeholders with shared goals might be the key for data sharing. Further, participant (E2) at Philips considered the access to right data a key component of experimentation with data, and therefore a major success factor for the New Product Development.

The second category is concerning the access to data by the organization's employees. Participant (A) from Algoritmica, and (B) at Datafloq stressed the important of having data centralized in order to provide the ability to all interested actors within the company to access and process it. In the theoretical framework, we briefly explained some of the strategies regarding Information Systems integration. We believe that in order to organize and manage successfully Big Data, most organizations should have a centralized approach. Further, we recognize the importance of acquiring data, therefore we would suggest organizations to establish "data-alliances" with partners, suppliers and other actors with common interests.

Extracting meaningful output and understanding the Value

Organizations should be aware that output from data is useful and efficient only if it encloses value. Morabito (2015), and Ohlhorst (2013) suggest that collecting data does not provide any value unless crucial information are extracted. In order to classify as important information, the result must be valid and accurate (Ohlhorst, 2013). Our participants suggested that understanding the value of data is one of the initial steps of success. Participant (A) at Algoritmica, and (C) at Thales, argued that it is vital to build and understand a business case with data, by combining managers and data scientists. Further, interviewee (E1) at Philips pointed out the importance by suggesting specialized data scientists and managers to examine the data offered and determine its potential. We understand the importance of extracting value information but we also understand the nature of new discoveries. In some cases in order to acquire a valuable output, it is necessary to experiment and reveal new methods, processes and results.

Legal and Security

A significant limitation of using Big Data which is not affected by the organization, is within the privacy and legislation context. Manyika et al. (2011) and Ohlhorst (2013), emphasize the policies regarding medical and financial records. In some countries, the limitation is includes more categories of private data protected by laws. Further, there have been incidents in the past of organizations with security breaches and the loss of important private information. Morabito (2015) explains that in some cases, the organizations responsible for securing the data were hiding the truth in fear of their public, corporate image.

Interviewee (E1) at Philips highlighted the need to have a legislation reform that will allows the good use of Big Data and benefit both organizations and the society by promoting innovation. Furthermore, participant (D2) at TNO agrees with the legislation change and adds that organizations should encourage the right mental model for the acceptance of data innovation. In order to achieve that, they should pick the right timing and the right framing. We believe that it is time for organizations and the society to start a constructive dialogue regarding Big Data and privacy. Companies should approach the government and the society and agree on a reform which will be accepted by all interested groups.

5.5 Key Success Factors

This part of the analysis is examining and comparing the theoretical framework as structured by the literature, and the empirical findings of the research. Participants were asked to identify which are – if any – the key success factors for successfully implementing Big Data analytics into the New Product Development process. Therefore, in order to determine the specific area of impact of each factor, the classification of the Key Success Factors was grouped in eight categories: Intent, Culture, Technology, Staff, Data, Process, Organizational, and Privacy.

The factor of intent is referring to the organization's actions to design and implement a Big Data strategy by choosing the suitable technologies and techniques according to their requirements (Manyika et al., 2011; Sathi, 2013). Participant (D1) at TNO suggested that it is essential for organizations to have set their objectives as well as a Data strategy in order to successful manage Big Data. Further, (D1) believes that a strategic collaboration with different stakeholders with common goals and ambitions, could unlock data sharing opportunities. Additionally, interviewee (C) at Thales explained organizations should oversee the interests of stakeholders in the data, rather than focusing only on handling data.

All participating organizations referred to the significance of a data-driven culture and the impact upon the connected operations. Parmar et al. (2014), explain that in a data-driven culture, data holds a central function and promotes a fact-based decision-making process, while data-driven decisions are more informed and effective (McAfee & Brynjolfsson, 2012). Interviewee (A) from Algoritmica, argued that a data-driven culture is more important than Big Data technology, since this is where limitations often occur. Participant (B) at Datafloq, added that a culture to involve data in every process could improve functions such decision-making. Furthermore, contributor (D2) from TNO suggested that organizations using data should encourage the acceptance of Big Data in the New Product Development process. In order to do so, organizations should promote a mental model for data innovation and analysis. Contributor (E2) at Philips connected the data-driven culture with the openness of experimenting and exploring new things with the use of data, as an important process for New Product Development. Finally, interviewee (C) at Thales emphasized the importance of data-driven culture and explained that organizations should understand that handling data has both direct and indirect influences.

According to the theoretical framework, the technological adoption is one of the significant factors for an effective commercial use of Big Data, especially regarding real-time decision-making (Morabito, 2015; Sathi, 2013; Ohlhorst, 2013). Despite this reference in the literature, only one participant emphasized the importance of technology as a key success factor. Specifically, interviewee (E1) at Philips and (F) at Apollo suggested that in several cases it is crucial to analyze the data and draw conclusions as fast as possible, in order to exploit possible market opportunities.

On the contrary, many contributors described the factor of having a well-trained staff as necessary to successfully organize Big Data. Manyika et al. (2011), includes the urgency of acquiring skilled human resources and believes that it will be a decisive factor that it will become a challenge due to a shortage. Participant (B) from Datafloq argued that there is insufficient specialized human resources, and believes that organizations might need to start training their own staff. In addition, interviewee (C) at Thales suggested that teams should be multidisciplinary and composed by specialists. Generalists-engineers should oversee the entire scope, operationally, tactically, and politically to ensure that the organization’s strategy is followed. Contributors (E1) and (E2) from Philips, pointed out that the staff working with Big Data, should have the necessary competencies, both in technical and commercial aspects, in order to achieve the goals set. Another success factor was highlighted by interviewee (F) which stretched the importance to have sufficient human resources but also have early adopters within the organization.

The theoretical framework of this thesis, described how data have changed the decision-making process and unlocked new business opportunities (DiFilippo & Blase, 2014; Morabito, 2014; Ohlhorst, 2013). Wengler (2001) further argues that having access to data enables the emerging of patterns and trends. Interviewee (B) at Datafloq believes that expanding data collection to external sources will improve insight gaining and promote data experiments. Participant (E2) from Philips suggested a similar approach, obtaining the right data and experimenting with them in order to explore new opportunities and create a sense of urgency. Finally, interviewee (D1) from TNO added that collaboration with different stakeholders with combined ambitions and goals will set the base for strategic data sharing.

The process of organizing the information systems and the data inside them is determined by the data-maturity of an organization in accordance with the strategy (Laudon & Laudon, 2011). Participant (A) from Algoritmica, highlighted that having centralized data which is accessible to everyone is a key factor of success, and could significantly promote project management and several other processes. Interviewee (B) at Datafloq described a similar approach of centralized and accessible data, which could enable a combination of data.

Manyika et al. (2011), argues that organizations need to change, adapt to Big Data requirements, while they should perceive and design the business process of implementing a data strategy. Contributor (A) from Algoritmica, suggested that organizations should build and understand the business case when managing data. Therefore, it is vital to promote the connection between data scientists and managers for an effective collaboration. Interviewee (D1) further added the value of speed in hyper-connectivity within the organization regarding the use of data and achieving the goals.

Two participants stressed the necessity for data-driven organizations to act in order to promote legislations enabling Big Data activities, and cope with that limitation. According to McKinsey (2011) and Kerr & Earle (2013), organizations and societies will have to position themselves and choose between privacy and efficiency in several market segments currently unavailable due to privacy laws. Interviewee (E1) from Philips urged organizations to ensure that future legislations will not prohibit the good use of Big Data that will benefit both societies and companies. Further, participant (D2) at TNO suggested that privacy laws must be addressed in order to allow the expansion of data-driven innovation, but in order to materialize organizations should choose the right framing and timing.

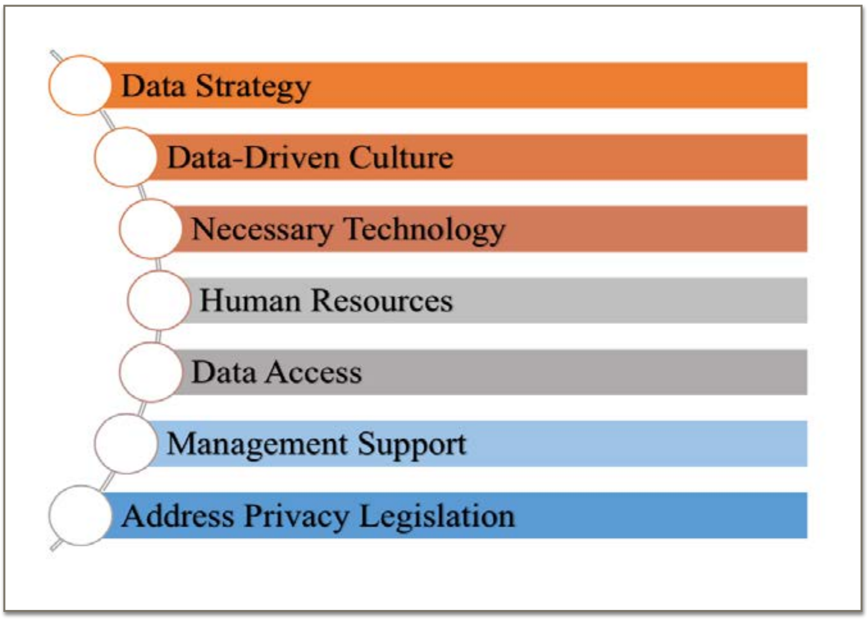


Exhibit 9. Summary of Key Success Factors

6. Conclusion

The objective of this study was to explore and provide an understanding of what role Big Data holds in the New Product Development and how companies in the Netherlands organize for Big Data. Furthermore, the study aimed to examine perceived benefits and managerial challenges linked with Big Data in the New Product Development.

In order to answer the main research question of this study: "*What is the role of Big Data in the New Product Development?*" it was essential to start by exploring how the theory and participants at chosen case companies identify Big Data and its functions. Undoubtedly, both the theory and the empirical findings illustrated inconsistency, and presented diverse definitions without any ties to industry, sector, company, department or function. However, in many cases Big Data was associated with recently emerged technological or managerial problems, while in others it was directed towards immense business opportunities. However, a clear pattern was disclosed. In the cases where the definition of Big Data was associated with limitations there was no clear organizational setup nor a Big Data objective. Contrary, in cases where Big Data was related to opportunities in development, production and enhancements, there was a clear approach with a well-developed business case that incorporated Big Data in its process to reach the goal.

Big Data enables the possibility for organizations to gain significant benefits, which can result into a competitive advantage. Benefits extend to a large spectrum of business activities can be translated into cost efficiency, insights or predictions, but also act as the key for New Business Development possibilities. Also, advanced tools offer the opportunity for organizations to manage large datasets at some extent, similar to economies of scale. Therefore, the more data the organization acquires, the cheaper it will become to store and analyze per data unit. With the increased data volume, organizations are able to expand the possibilities of market research and customer targeting more effectively, and support the New Product Development. A less expected outcome of Big Data is caused by its disruptive action, and the ability to unlock and compete in new non-core markets that otherwise would be impossible without the role of Big Data.

The process of using collected data to extract insights and value is not new, but due to the exponential growth of data volume and its variety, organizations are required to implement new tools and techniques. The empirical findings conclude that the role of Big Data is widespread within the NPD, however its influence is more likely to be aimed on supporting already employed NPD structures than enabling novel ones. Despite the lack of knowledge and initiative to set up new NPD structures and processes which incorporate Big Data as a standard element, Big Data holds a crucial but indirect role in the established NPD processes. For the objective of this study the two main models of NPD, the Stage-Gate model and the Lean Startup method, were investigated which also proved to be mostly employed by the case companies.

The empirical findings show that the variations of Stage-Gate model are undoubtedly still the most common approach to NPD in more traditional industries or product categories especially that require big investments and high-degree of formalization. The role of Big Data in the pre-development (NCD model) stage was found to strengthen important activity elements and segments. The effect of the influencing factors streaming in to the NCD model, from external and internal sources, increase with Big Data due to informationalization. However, the empirics as well as the theory (Keon et al., 2001) highlight that the embedded corporate culture, governance and strategies are directly influencing the role of Big Data within the process. Nonetheless, the circular NCD model is more flexible than more sequential processes, thus better suited to adopt Big Data into its process of generating concepts. In novel models, such as the Lean Startup model, data-enabled feedback loops and insights are a critical component of success. Flexible NPD models can draw more benefits from data-driven approaches, because of their dynamic nature of swift implementation of new changes. Also, in order to enable an effective operation of models and enhance the value generation throughout the chain, organizations should collect data according to their needs. Some cases require all the possible available data, while others demand a targeted acquisition of datasets. In addition, organizations have to ensure the high quality of the data before they base their business and operational decisions on it. Further, to stimulate NPD through Big Data there is a need for alignment between Big Data strategies and overall business strategies to ensure effectiveness. In other words, top management has to continuously provide clear guidelines and present an enterprise-wide business case that can effectively be solved with the means of Big Data.

Further, the successful implementation of Big Data requires organizations to be prepared to address some - significant for succeeding - challenges. Apart from solving technical challenges, it is vital for data-driven organizations to focus on managing the managerial dimension of these obstacles. The first limitation to overcome is connected with securing the resources (human and financial) required for the transition into a data-driven organization. But perhaps the most important step is the design and implementation of a data-driven culture, which requires employees to understand the important role of data as well as have a centralized and clear access to it.

6.1 Recommendations

This chapter will provide recommendations based on the theory and empirical findings. The following steps are to provide an initial road-map for organizing Big Data within the organizational functions, with an emphasis in the New Product Development.

- ❖ Urge the Top-Management to create a Data Strategy according to the needs of the organization that will enable employees to access data through centralized platforms.

- ❖ Establish a Data-driven Culture which will involve as many employees as possible, since evidence-based decisions are more effective decisions.

- ❖ Appoint a Chief Data Officer with the responsibility of connecting managers and data scientists, as a formula to bridge the data-understanding gap between these two groups.

- ❖ Train existing employees or create multidisciplinary teams in order to address the data scientist shortage.

- ❖ Promote Data-Alliances with stakeholders which share common goals, and gain access to valuable sources of data.

- ❖ Adoption of flexible forms of New Product Development models, since they can benefit more from Big Data utilization compared to traditional approaches.

- ❖ Emphasize Data-enhanced feedback loops, to add product improvements or redesign the products.

- ❖ Encouragement of data experimentation, as a mean to promote Data-driven Innovation and Product Champions, in the New Product Development.

- Campaigns and constructive dialogues with the society and state institutions, with the goal of reaching an agreement regarding the reformation of privacy legislation.

6.2 Future Research

As we mentioned earlier in the thesis, we believe that the literature describing the role of Big Data in the New Product Development is inadequate. Therefore, our goal was to examine the role of Big Data in multiple functions of the New Product Development stages. However, the focus was undulated between a diverse number of industries, since our intention was to create a broad comprehension on how Big Data is applied in different New Product Development processes. In a future research, this exploratory approach could be replaced by an in-depth examination of specific industries that are considered successful in Big Data implementation in the New Product Development. This investigation would provide interesting results regarding the frameworks, methods and factors that stimulated their success, from both the managerial and operational point of view.

The empirical findings of this study indicated that the most successful Big Data applications are connected to predictive maintenance through real-time data generating devices with sensors. Despite that, we conceptualize significant undiscovered potential in the Internet of Things and Smart-Products, an interactive category of products which is still in its infancy. Further, smart-products are constantly reporting back all kinds of data that synthesize useful information and they are currently limited by privacy and security reasons, and the lack of know-how to process them effectively. Organizations are currently experimenting on how the Internet of Things can affect Smart-Products. From a researcher’s point of view, examining how to organize the New Product Development of Smart-Products, by using the Internet of Things would compose an interesting and challenging topic. However, we believe that the best timing to engage to such research would be when organizations will establish a higher degree of maturity regarding Big Data implementation and the Internet of Things.

All the above advocate that Big Data can be an important component of new robust products, or offer product improvements sewed accordingly to their users’ needs. Furthermore, considering the fact that Europe together with the rest of the Western economies are in the process of becoming wide-service providers, it would be interesting to examine how Big Data can enhance the New Service Development process. In order to achieve that, some of the dimensions related to product development should be replaced to match the new research purpose.

7. References

Afuah, A. (2002). *Mapping technological capabilities into product markets and competitive advantage: the case of cholesterol drugs.* Strategic Management Journal. Vol 23(2): p. 171-179. (2002).

Alam, I. (2006). *Removing the fuzziness from the fuzzy front-end of service innovations through customer interactions.* Industrial Marketing Management, Vol. 35(4): p. 468-480.

Bacon, G., Beckman, S., Mowery, D. & Wilson, E. (1994). *Managing Product Definition in High-technology Industries:A Pilot Study*. California Management Review, 36(3), 32- 56

Barlow, M. (2013). *The Culture of Big Data.* Sebastopol, CA: O'Reilly Media, Inc.

Belliveau, P., Griffin, A., & Somermeyer, S. (2002). *The PDMA ToolBook for New Product Development.* New York, NY: John Wiley & Sons, Inc.

Bessant, J., Lamming, R., Noke, H. & Phillips, W. (2005). *Managing Innovation Beyond the Steady State.* Technovation. Vol. 25(12): p. 1366-1376.

Boeddrich, H-J. (2004). *Ideas in the Workplace: A New Approach Towards Organizing the Fuzzy Front End of the Innovation Process*. Creativity and Innovation Management. Vol. 13(4): p. 274-285.

Boer, H. & During, W.E. (2001). *Innovation, What Innovation? A Comparison between Product, Process and Organisational Innovation.* International Journal of Technology Management. Vol. 22. Nos 1/2/3.

Bryman, A. & Bell, E. (2011). *Business Research Methods.* 3rd Edition, Oxford University Press, USA.

Burrell, G. & Morgan, G. (1979). *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life.* London, UK: Heinemann Educational Books.

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

Chesbrough, H., Vanhaverbeke, W. & West, J. (2006). *Open Innovation: Researching a New Paradigm.* New York, NY: Oxford University Press Inc.

Cooper, R.G. (1998). *Benchmarking New Product Performance: Result of the Best Practices Study".* European Management Journal. Vol. 16: p. 1-17.

Cooper, R.G. (2001). *Winning at New Products: Accelerating the Process from Idea to Launch.* 3rd Edition. Cambridge, MA: Perseus Publishing.

Cooper. R.G. (2008). *Perspective: The Stage-Gate Idea-to-Launch Process - Update, What's New, and NexGen Systems.* Journal of Product Innovation Management. 2008;25: p. 213-232. (2008).

Cooper. R.G. & Edgett, S.J. (2010). *Developing a Product Innovation and Technology Strategy for Your Business.* Research Technology Management. Vol. 53: p. 33-40.

Damanpour, F. & Schneider, M. (2006). *Phases of the Adoption of Innovation in Organizations: Effects of Environment, Organisation and Top Managers.* British Journal of Management, Vol. 17: p. 215-236. (2006).

Davenport, T.H. (2014). *Big Data at Work.* Boston, MA: Harvard Business School Publishing.

Davila, T., Epstein, M.J. & Shelton, R.D. (2013). *Making Innovation Work.* Updated edition. Upper Saddle River, NJ: Pearson Education, Inc.

Dodgson, M., Gann, D. & Salter, A. (2008). *The Management of Technological Innovation.* New York, NY: Oxford University Press Inc.

Dumbill, E. (2013). *Making Sense of Big Data.* Big Data. 1(1); p.1-2. (2013).

Feinleib, D. (2014). Big Data Bootcamp: *What Managers Need to Know to Profit from the Big Data Revolution*. New York, NY: Apress Media Inc.

Galbraith, J.R. (2014). *Organization Design Challenges Resulting from Big Data*. Journal of Organization Design. Vol. 3(1): p. 2-13. (2014).

Garcia, R. & Calantone, R. (2002). *A Critical Look at Technological Innovation Typology and Innovativeness Terminology: A Litterature Review.* Journal of Product Innovation Management. 2002;19: p. 110-132. (2002).

Gassmann, O., Sandmeier, P. & Wecht, C.H. (2006). *Extreme Customer Innovation in the Front-end: Learning From a New Software Paradigm*. International Journal of Technology Management. Vol. 33(1): p. 46-66.

Gobble, M. M. (2013). *Big Data: The Next Big Thing in Innovation.* Research Technology Management. Vol. 56(1) January - February (2013). Industrial Research Institute, Inc.

Gopalakrishnan, S. & Damanpour, F. (1996). *A Review of Innovation Research in Economics, Sociology and Technology Management.* Omega. Vol. 25(1): p. 15-28. (1997).

Grant, R.M. (2010). *Contemporary Strategy Analysis.* 7th Edition. West Sussex, UK: John Wiley & Sons Ltd.

Grossman, R.L. & Siegel K.P. (2014). *Organizational Models for Big Data and Analytics*. Journal of Organization Design. Vol. 3(1): p. 20-25. (2014).

Gubbi, J., Buyya, R., Marusic, S. & Palaniswami, M. (2013). *Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions.* Future Generation Computer Systems. Vol. 29(7): p. 1645-1660. (2013).

Heller, T. (2000). *If Only We'd Known Sooner: Developing Knowledge of Organizational Changes Earlier in the Product Development Process.* IEEE Transactions on Engineering Management. Vol. 47(3): p. 335-359.

Herstatt, C. & Verworn, B. (2001). *The "Fuzzy Front End" of Innovation.* Working Paper. Department for Technology and Innovation Management, Technical University of Hamburg.

Kerr, I. & Earle, J. (2013). *Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy.* Stanford Law Review. Vol. 66(65).

Kim, J. & Wilemon, D. (2002). *Focusing the Fuzzy Front-end in New Product Development.* R&D Management. 32(4). (2002)

Koen, P., Ajamian, G., Burkart, R., Clamen, A., Davidson, J., D'Amore, R., Elkins, C., Herald, K., Incorvia, M., Johnson, A., Karol, R., Seibert, R., Slavejkov, A., & Wagner, K. (2001). *Providing Clarity and a Common Language to the "Fuzzy Front End"*. Industrial Research Institute, Inc.

Kohn, K. (2005). *Idea Generation in New Product Development Through Business Environmental Scanning: The Case of XCar*. Marketing Intelligence & Planning. Vol. 23(6/7): p. 688-704.

Kopetz, H. (2011). *Real-time Systems: Design Principles for Distributed Embedded Applications.* 2nd Edition. New York, NY: Springer Science+Business Media, LLC.

Khurana, A. & Rosenthal, S.R.(1998). *Towards Holistic "Front Ends" in New Product Development*. Journal of Product Innovation Management. Vol. 15(1): p. 57-74.

Lake, P. & Drake, R. (2014). *Information Systems Management in the Big Data Era*. Switzerland: Springer International Publishing.

Langerak, F., Hultink, E.J. & Robben, H.S.J. (2004). *The Role of Predevelopment Activities in the Relationship Between Market Orientation and Performance*. R & D Management. Vol. 34(3): p. 295-309.

Laudon K. & Laudon J. (2011). *Management Information Systems: Managing the Digital Firm*. Upper Saddle River, NJ: Pearson Education, Inc.

Lotz, M., Brent, A.C. & Steyn, H. (2009). *Investigating the Risk Management Potential of a Stage/Phase-Gate Project Management Approach.* Journal of Contemporary Management. Vol. 6: p. 253-273.

McAfee, A. & Brynjolfsson, E. (2012). *Big Data: The Management Revolution.* Harvard Business Review 90(10): p. 60-68.

Moenaert, R.K., De Meyer, A., Souder, W.E. & Deschoolmeester, D. (1995). *R&D/Marketing Communication during the Fuzzy Front-end.* IEEE Transactions on Engineering Management. Vol. 42(3): p. 243-258.

Morabito, V. (2015). *Big Data and Analytics: Strategic and Organizational Impacts.* Switzerland: Springer International Publishing.

Morabito, V. (2014). *Trends and Challenges in Digital Business Innovation.* Switzerland: Springer International Publishing.

Nagji, B. & Tuff, G. (2012). *Managing Your Innovation Portfolio.* Harvard Business Review 90(5): p.66-74.

Nasierowski, W. (2008). *Project Development: Why does it really matter to companies?.* Institute of Organisation and Management in Industry, Vol. 1(1): p. 63-68. (2008).

Nidumolu, R., Prahalad, C.K. & Rangaswami, M.R. (2009). *Why Sustainability is Now the Key Driver of Innovation.* Harvard Business Review 87(9): p. 57-64.

O’Connor, G.C. (1998). *Market Learning and Radical Innovation: A Cross Case Comparison of Eight Radical Innovation Projects.* Journal of Product Innovation Management. 1998;15: p. 151-166.

Ohlhorst, F. (2013). *Big Data Analytics: Turning Big Data into Big Money.* Hoboken, NJ: John Wiley & Sons, Inc.

Pellissier, R. (2008). *A Conceptual Framework for the Alignment of Innovation and Techhnology*. Journal of Technology Management & Innovation, Vol. 3(3). (2008).

Robbins, P. & O’Gorman, C. (2014). *Innovating the Innovation Process: An Organizational Experiment in Global Pharma Pursuing Radical Innovation.* RADMA and John Wiley & Sons Ltd. Vol. 45(1): p. 253-273. (2015).

Rogers, E.M. (2003). *Diffusion of Innovations.* 5th Edition. New York, NY: Simon & Schuster, Inc.

Sathi, A. (2012). *Big Data Analytics: Disruptive Technologies for Changing the Game*. IBM Corporation. Boise, ID: MC Press.

Schaltegger, S. & Wagner, M. (2010). *Sustainable Entrepreneurship and Sustainability Innovation: Categories and Interactions.* Business Strategy and the Environment, 20: p. 222-237. (2011).

Simon, P. (2013). *Too Big to Ignore: The Business Case for Big Data.* Hoboken, NJ: John Wiley & Sons, Inc.

Stevens, E. (2014). *Fuzzy front-end learning strategies: Exploration of a high-tech company*. Technovation, 34: p. 431-440. (2014).

Tidd, J. & Bessant J. (2009). *Managing Innovation: Integrating Technological, Market and Organizational Change.* 4th Edition. Chichester, UK: John Wiley & Sons Ltd.

Trott, P. (2011). *Innovation Management and New Product Development.* 3rd Edition. Essex, UK: Pearson Education.

Tuomi, I. (2002). *Networks of Innovation: Change and Meaning in the Age of the Internet.* New York, NY: Oxford University Press Inc.

Van den Bosch, J. & Duysters, G. (2014), *Corporate Venturing: Organizing for Innovation*. Cheltenham, UK: Edward Elgar Publishing Limited.

Vare, T. & Mattioli, M. (2014). *Big Business, Big Government, Big Legal Questions.* Managing Intellectual Property. (2014).

Wengler, J. (2001). *Managing Energy Risk: a Nontechnical guide to Markets and Trading.* Tulsa, OK: PennWell Publishing Company

Wolfe, R. A. (1994). *Organizational Innovation: Review, Critique and Suggested Research Directions.* Journal of Management Studies. Vol. 31(3): p. 405-431. (1994).

Xia, F., Yang, L.T., Wang, L. & Vinel, A. (2012). *Internet of Things.* International Journal of Communication Systems. Vol.25: p. 1101-1102.

Yin, R. K. (2011). *Qualitative Research from Start to Finish*. New York: Guilford Press.

Electronic references:

Blank, S. (2013). Why the Lean Start-Up Changes Everything. Harvard Business Review. Retrieved from: https://hbr.org/2013/05/why-the-lean-start-up-changes-everything [Accessed: 21/04/2015]

Bloem, J., van Doorn, M., Duivestein, S., van Manen, T., van Ommeren, E. & Sachdeva, F. (2013). *Your Big Data Potential: The Art of the Possible*. VINT. Retrieved from: https://us.sogeti.com/wp-content/uploads/2014/04/Big-data-report-4.pdf. [Accessed: 14/04/2015].

Capgemini. (2015). *Big & Fast Data: The Rise of Insight-Driven Business*. Retrieved from: https://www.capgemini.com/resource-file-access/resource/pdf/big_fast_data_the_rise_of_insight-driven_business-report.pdf [Accessed:17/03/2015].

Capgemini. (2015). *Cracking the Data Conundrum: How Successful Companies make Big Data Operational*. Retrieved from: https://www.capgemini-consulting.com/resource-file-access/resource/pdf/big_data_pov_03-02-15.pdf. [Accessed: 15/02/2015].

Cooper, R.G. & Kleinschmidt, E.J. (2001). *Stage-Gate Process for New Product Success*. Innovation Management U3. Retrieved from: http://wiki.bio.dtu.dk/teaching/images/3/3c/Stage-gate.pdf [Accessed: 29/04/2015]

DiFilippo, D. & Blase, P. (2014). *Gut & Gigabytes: Capitalising on the Art & Science in Decision Making*. PricewaterhouseCoopers. Economist Intelligence Unit Report. Retrieved from: http://www.pwc.com/es_MX/mx/servicios-tecnologias-de-la-informacion/archivo/2014-10-big-decisions.pdf [Accessed: 29/03/2015]

European Commission. (2013). *Business Opportunities: Big Data*. Retrieved from: https://ec.europa.eu/growth/tools-databases/dem/sites/default/files/page-files/big_data_v1.1.pdf [Accessed: 07/04/2015]

Harris, J.G., Shetterley, N., Alter, A.E. & Schnell, K. (2013). *The Team Solution to the Data Scientist Shortage*. Retrieved from: http://www.accenture.com/sitecollectiondocuments/pdf/accenture-team-solution-data-scientist- shortage.pdf [Accessed: 14/4/2015]

Industrial Research Institute. (2004). *New Concept Development model: Front-End Innovation.* Retrieved from: http://web.stevens.edu/cce/NEW/optFEI.htm [Accessed: 04/03/2015]

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. & Byers, A.H. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute. Retrieved from: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation. [Accessed: 21/02/2015]

OECD. (2014). *Data-driven Innovation for Growth and Well-being: Interim Synthesis Report*. Retrieved from: http://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf. [Accessed: 07/04/2015]

Ozzimo, A. (2014). *Real-Time Versus 'Right-Time' Decision Making: A New Competitive Advantage*. Retrieved from: http://www.oracle.com/us/cio-docs/5-minutes-modern-decision-ar-2190457.pdf [Accessed: 27/03/2015]

Parmar, R., Cohn D.L. & Marshall, A. (2014). *Driving Innovation through Data*. IBM Institute for Business Value. Retrieved from: www-935.ibm.com/services/us/gbs/thoughtleadership/innovation-through-data/ [Accessed: 14/04/15]

Ross, J.W., Beath, C.M. & Quaadgras, A. (2013). *You May Not Need Big Data After All*. Harvard Business Review. Retrieved from: https://hbr.org/2013/12/you-may-not-need-big-data-after-all [Accessed: 12/03/2015]

SAS Institute. (2013). *Five Big Data Challenges*. Retrieved from: http://www.sas.com/resources/asset/five-big-data-challenges-article.pdf [Accessed: 11/03/2015]

Sebell, M.H. (2008). *Stage Gates: Good or Bad for Innovation?*. CreativeRealities, Inc. Retrieved from: http://www.creativerealities.com/portals/64975/docs/stagegatesgoodorbad.pdf [Accessed: 27/03/2015]

8. Appendixes

Appendix A. Empirical Findings in Text

A1. Big Data Definition

Participant (A) explains that data has been increasing exponentially in size and volume for a long time. Also, he pointed out the increase in the computing power and the need to use tools such as predictive models, anomaly detection and natural language processing, in order to solve more complicated problems, connected with the radical growth of data size. Interviewee (A) claims that too much attention is being directed towards the term “Big Data”, he argues that organizations should rather talk about just “data”. Participant (A) concludes that if data is an important component of the product it should be considered Big Data, regardless of its volume.

Participant (B), highlights that a more accurate term for Big Data is “Mixed data”. Further, he defines that Big Data is a combination of internal and external data sources to gain better understanding and insight. Interview Participant (C) defines Big Data as enormous amount of data that is opening up, being both internal and external. She refers to Big Data as an exponential and a more complex data world. Interviewee (F) defines Big Data as the amount of data that allows for complete coverage of all settings which can happen and enable useful data combinations. Further, (F) emphasizes that Big Data is such data that enables reliable predictions without knowing anything about the relationship between input and output.

Interviewee (D1) at TNO explains that the definition of Big Data is the least important question, and that there exist a lot of definitions but mostly definitions of technological problem. Big Data allows us to discover things in a larger scale that we could not discover in small scale. Further, do things with the data that was not possible earlier.

Participant (D2) believes that the best definition to describe Big Data is involving any data-related problem where the size of the data becomes part of the problem. Further, (D2) adds the tools, methods and approaches that enables us to do things with data which were not feasible in the past. The definition that (D2) is providing, implies that the problem of the Big Data size is changing together with the technological advancements on the data volume and the analytic tools. An illustration of his definition is reflected in the phrase that “things we call Big Data now, it will be in people’s phones in 10 years”.

The interviewee (E1) at Philips relates Big Data to the unstructured raw data obtained from production and a large set of applications, which is considered to be less accurate. Participant (E2), explains that for him Big Data is not the amount of data but rather the difficulty in understanding what is in the data. Therefore, when the difficulty becomes high, he suggests the term Big Data. Further, he characterizes Big Data as complex, ambiguous, unstructured data from multiple internal and external sources.

A2. Big Data Objective

According to Interviewee (A), the main objective of Big Data usage is to predict failure in mechanical devices and observe their performance. Data is used to improve the projects and ensure that parameters are set. Interviewee (A) highlights that the complexity of the objective determines the required skill set and analytical tools, not the amount of data itself. However, Big Data is a vital input to see if their own and their customers products function correctly. Further, he (A) illustrates the above with an example that before 2008 these devices did not have any data aggregation, and maintenance monitoring, especially the offshore, was costly. That acted as an incentive to add sensors in the mechanical devices, in order to remote monitor, identify malfunctions and avoid downtime. As a result, the company aims to develop a scalable software product by using data aggregation.

Participant (B) explains that the most common objective of using Big Data is to gain a greater understanding on how products and users behave, to be able to establish more accurate recommendations for new concepts of services or products.

Interviewee (C) explains that Thales want to achieve internal and external security goals from using Big Data. In many cases data is provided by other companies that have the need to secure it. Therefore, Thales' objective is to identify the potential and the need for data security, such as open data in sensitive industries which needs to be protected. By doing so data can be transformed into security products and solutions.

Interviewee (D1) explains that Big Data objectives are to provide insights to be able to transition from the more traditional way of looking and approaching the problem. These insights can translate into more personalized products and higher precision. Also, to develop new products and solutions, which are more predictive, contextualized and automated. Further, the interviewee (D1) emphasizes that Big Data objectives are to forge new value networks, and the impact of data is restructuring core segments and networks.

According to participant (D2), Big Data is a way for organizations to uncover patterns that are hidden, and use them for commercial or social reasons. In addition, (D2) believes that Big Data are providing insights and assist in making prediction on several different markets and sectors, despite the fact that prediction models are in their infancy. Finally, interviewee (D2) considers Big Data as a tool of confirmation on assumptions and decisions that organizations were already suspecting.

At Philips, Interviewee (E1) points out different Big Data objective due to the size of the company. To extend products and services, ensure consumer satisfaction, and improve product quality are determined to be Big Data objectives. Other objectives are connected to reduction of cost and higher efficiency in production, as well as discovery of new markets through data.

Interviewee (E2), describes some objectives set for the usage of Big Data. Specifically, Big Data is used for accurate prediction on sales, understand what qualities users value more and create something tailored to it. Furthermore, participant (E2) adds that data can be used for discovering and developing new innovations. In order to achieve that, Big Data can support with information acquisition and to discover trends. The results can suggest which products to create, what kind of technologies will be available and what are the prime demands for users.

Participant (F) explains that Apollo Vredestein lacks a formal or established Big Data objective at the moment. However, a general objective of having good products and bringing development costs down, thus Big Data could be a useful tool to do so.

A3. Data Management

Participant (A) considers that all available data is useful if it is in any way connected to identification of problems in the mechanical devices that the company's software is monitoring while other data is excluded. The sources of data are from sensors and maintenance logs containing CMS (Content Management System) data, SCADA (Supervisory Control and Data Acquisition) data, Oil pressure- and temperature data, and performance register data. Sensors send data periodically in near real-time to a central fleet database where the company's software is installed to register the data, and evaluate its utility and quality. Interviewee (A) explains that problems which arise are based on the nature of real data - where missing values and data have to be sorted with the use of robust techniques and scripts to ensure data quality. The company's data assessment starts with a proof of concept, where relevant problem samples or historical data and logs are provided in order to develop the solution of predictive analytics. Further, (A) argues that if Big Data projects are provided with all kinds of data in hope that useful insight and value can be extracted from it, they will most likely fail.

According to Interviewee (B), the usefulness of data depends on the analytic or business objective. He (B) explains that data quality is established using complete and accurate data which includes all variables and aspects relevant to the data mining objective. Besides from poor data quality, wrong application of it or wrong processing of it (incorrect algorithms) will lead to wrong insight and failure to extract value.

Interviewee (C), explains that Thales´ main data acquisition source is their own machine-generated data, primarily from sensors and radars. Further, Thales has access to external data in the forms of open data, such as social media, and data from customers with a commercial value. Regarding the quality of the data, participant (C) states that their own machine-generated data possess a high quality, therefore they prioritize it over data from unreliable sources. Even if the quality may be good, Thales´ data experts perceive it as non-good or unimportant if it derives from untrusted sources.

According to Interview Participant (D1) at TNO, the gathering and usefulness of data needs to be based on the objective and the purpose of the project. We need to consider if the data set is suitable for the right questions to be answered. Also, the quality of the data can be sufficient but not for the purpose the organization wants to use it for. Further, interviewee (D1) explains that the quality of the data is connected to the limitation of data and the anticipated results. Participant (D1) explains that Linked data allows to combine different data sets in an open format. Furthermore, metadata can include information about the quality of the data set which is beneficial for the user to access.

Interviewee (D2), suggests that data collection should be a combination of data already possessed and data you have not thought of in the past. Also, organizations should form agreements and collect open external data from collaborators. Regarding to which data is considered useful, he argues that there are two different schools of thought. The first school wants all the data they can get, and the second school that requires only specific sets of targeted data that the organization is seeking and knows how and where to apply them. Further, (D2) argues that the quality of the data depends on the purpose of usage and the requirements. Furthermore, he believes that there are three kinds of useful data, production data, data influencing the development process, and additional undiscovered data that no one thought they would an effect or value.

Interviewee (E1) explains that Philips does not have a common approach on how to capture and collect data. He (E1) describes that production data, external data, and to some extent "usage" data from their products is being utilized. Privacy laws are causing difficulties for data aggregation from the field/device data, such data is considered as highly useful and valuable for Philips according to Interviewee (E1). Philips use a wide set of tools to analyze and assess data, which tools are used depends on the purpose and objective of the project. According to Interviewee (E1), there is a need of good quality data and trained staff to process and understand it. Participant (E1) emphasizes that regardless of purpose, the organization dealing with Big Data needs to hold an understanding of the acquired data, its background and the objective, to be able to successfully interpret it and make use of it.

According to interview participant (E2), there are both internal and external sources of acquiring data. The primary sources are the own product-generated data internally or externally, and public data, such as social media data. In addition, brokering data occurs but selectively. Further, he describes that it is important to anonymize the data and comply with the company's values and with the privacy regulations. Therefore, data which violates privacy and regulations or is unreliable is reject or deleted. Regarding data quality, interviewee (E2) claims that a way to determine the reliability of a data source is by the accuracy of past data-based predictions. Finally, using sufficient correlation in the production enables the ability to find significant additions to quality and improve accuracy.

Interviewee (F) at Apollo Vredestein argue that collecting data they can learn how their products are being used and store such data in a systematic way to create valuable relations. He (F) explains that collecting manufacturing data the company is able to link input parameter data to the output. Technological advancements have further enabled potential, such as having sensors in tyres to receive historical data which is seen as an access to a big source to understand the product better and enhance product development. When it comes to other types of data such as sales data, he (F) explains that the tyre industry is a closed and small environment which makes it difficult to collect Big Data to combine. However, there lies a lot of potential of accessing and combining data with other industries that are related to the own product.

Participant (F) explains that there is no clear approach at Apollo Vredestein on ensuring data quality. Assessment is often made after the initial relations are made by removing errors or false datasets to make relations more reliable. He (F) specifies that verifying quality in the product development process is easier since the data is cleaner than in production or manufacturing. Often it is one user that requests certain tests and verifies a specification in a loop which ensures quality.

A4. Organizational Setup

According to participant (A), the Big Data trend has increased top management's awareness of the potential value that data can bring into business development and operations. Companies have large volumes of data but it is scattered around multiple departments and databases, consequently killing data applications. Combining data can create a lot of promising applications. However, many companies do not have a central data warehouse or a fleet database to merge data, which would make it easier to work with data, combine the data and perform quality controls. Further, participant (A) mentions that with so much data available, everyone should have skills in data analysis to some extent, since decisions-making should be based on information. Therefore, it is important that everyone has access to data and not just the IT department. Successful companies, centralize the data location, and provide full and fast accessibility in order to have more informed decisions in all levels of the organizations.

Interviewee (B) argues that collecting and sharing of data needs to be promoted, therefore the organizational setup has to be structured for the purpose of enabling such features. Single departments should have access to most data to establish a higher data-driven culture. The founder of Datafloq (B) further explains that successful organizations designate and have a Chief Data Officer (CDO), acting as a bridge between business executives and data technology. The senior management's involvement is critical since for being successful, Big Data incentives have to take a top-down approach.

Interview participant (C) argues that Big Data has resulted in big organizational changes. The modern setting is not only to develop products or solutions for selling and maintaining them, there is a greater need for having an application of the product or solution and its technology. She reasons that Big Data impacts the organizations by putting higher importance on offering consultancy services together with the products. To implement Big Data it requires the right people, right methods, and right culture. Further, she explains that top management's involvement in Thales is important and evident for the implementation. Leading actions towards usage of data in cyber environments (e.g. 24/7 monitoring center) are boosted by the top management so that Thales can proactively advise on threats and offer products and solutions.

Interviewee (D1) explains that globalized platforms should be connected with decentralized problems and areas. He (D1) further argues that Big Data technologies are becoming more democratized. The organization has to encourage accessible data by using visualization techniques and sandbox environments so that non-data staff (individuals without data expertise) can be allowed to work with the data. Participant (D1) reasons that people with domain expertise have to be able to ask questions that can be answered with data, and understand how to bridge that gap.

According to participant (D2), organizations should compose a structure of flexible teams with the ability to test data, experiment with data, and validate data. He (D2), further argues that teams with some autonomy that are allowed to break some process regulations or even fail, have the potential to deliver useful results. Finally, (D2) suggests that data scientists and business managers should change their thinking focus on common goals and achievements.

Participant (E1) at Philips describes that Market Intelligence units analyze data from the field to capture trends and feelings which will be considered throughout NPD. Further, every department within the organization have specific setups for using Big Data. Interviewee (E1) cannot see any large and formal initiatives or development of functions to promote Big Data usage within Philips. However, there are huge efforts in terms of Digital Connectivity within Philips. Those initiatives are to make sure that the data flow from Philips to outside world and back is being done properly and in a good use. Interviewee (E1) further explain that Philips is adding an important element to the Architectural Function (architect composing and decomposing the product into its functions), which is to get Big Data from the field and translate it into certain product requirements.

Participant (E2), believes that in order to organize Big Data, organizations should have a core entity managing data from sales and the market. Thus, it is important to install a department focusing on Big Data. Additional, in order to manage research and innovation it is necessary to have new approaches with data. Teams with application knowledge that are focusing on certain markets and collect specific data, are effective and should be promoted. Finally, data scientists and experts should be spread all over the organization, in order to enable competencies.

Interview participant (F) explains that there is no general setup implemented to manage Big Data, and no fleet of data scientist working with Big Data. The company's data initiatives are solely based on a project basis with different objectives. However, the company has established electronics groups (to establish automated analysis systems) that have experience with Big Data and will support and implement the usage of Big Data analytics within Apollo Vredestein.

A5. New Product Development Structure

Interviewee (A) points out that they are not using an enterprise or textbook model at the basis of how to approach the New Product Development. He stretches the importance of being lean in the NPD process, since any advanced models will delay and set them back. Therefore, they need to prioritize rigorously by using simple tools. He describes their process as early staged, with customer validation, Kaizen method as well as using revision control and other project management tools. Interviewee (A), further explains that their NPD process is to first develop the basic features, shift the product as early as possible to the customer for validation, to receive feedback.

Participant (B) has pointed out some tools and methods to stimulate the New Product Development process. First, he described how a simulation analysis can strengthen concepts and features, as well as how it acts as an incentive to determine what you are looking for. He (B) also mentioned that ideas are generated from customer perspectives, and by market research and behavior analysis.

Participant (C), describes Thales´ NPD structure by referring to two applied models, depending on the type of product. Some business segments require a strict Stage-Gate model, testing and validating through Go/No-Go criteria. However, the Stage-Gate model is not used for all innovations, the needed flexibility for other products and solutions leads to the adoption of the

Lean-Startup model. Security-related development process requires to be flexible since continuous adjustments need to be allowed throughout the process. Furthermore, disruptive systems are needed in NPD for discovering and entering new markets and value networks. She argues that NPD for the market is about listening to the market, while for technology it is stricter. Borders of the NPD are becoming vague with the open innovation approach. Interviewee (C) describes Thales´ FFE phase as a concept bank, where ideas from internal and external sources, market analysis and objectives from groups are seen as feeding streams. In addition, FFE activities can be initiated both from a bottom-up approach (organic growth), where a group of individuals generates and defines a concept to build on, and from a top-down approach, with more directives for certain accomplishments which can be further supported by M&As. She explains that all developments have to in some way contribute to Thales, being financial, brand or image contributions.

Interviewee (D1) reasons that data-driven NPD is not very different in structure from other traditional NPD. However, the NPD needs to have an iterative approach especially for novelty. He (D1) explains that a lot of core data technologies (sensors, clouds etc.) intersect with vertical products and demands. Companies are starting to bend horizontal technologies into vertical solutions. Participant (D1) explains that the alternative to development, especially in data-driven innovation, is acquisitions of start-ups and solutions. This is due to the fact that companies notice that the landscape is changing but do not possess right capabilities on their own to adapt.

Participant (D2), explains that the New Product Development life-cycles involve a funneling approach. Several ideas are filtered through a stage-gate process, and details are added progressively to products. Additionally, (D2) adds that data-driven innovation and Big Data application development aren’t that different from traditional NPD processes in in their structure. He (D2), observes some common characteristics between the data-driven approach, the startup approach (market release as soon as possible), and also the open innovation approach.

Interviewee (E1) describes the NPD process in Philips as a Stage-Gate model process - with the initial phase (proof of assumptions), reference architecture phase (creation of design), and realization phase (launch to market). In the initial phase, a team, which its size depends on the scale of the project, develops a concept or solution to test in the market (product, service or combinations). Interviewee (E1) describes that the initial phase is incorporating a lot of theory and consumer group testing, to observe if targeted consumers are enthusiastic and notify the concept proposal. Another important part of the NPD is technical feasibility - to prove that the technical solutions are sufficient for mass production. He (E1) points out that the structure of the FFE varies based on the type and size of the product. Bigger projects and products receive more resources. However, the mechanism and principal is always the same, there is an assumption or an idea that will be validated. Validation involves experts and experienced individuals within different aspects of the product area. According to Participant (E1), after validation the team tries to implement and see if they can control the cost. Depending on the product it can take from 6 months to 3 years (e.g. large imaging systems) before the concept proposal is validated and forwarded to a reference architecture (key point of formal development). Interviewee (E1) states that during the formal development Philips have access to a lot of tools to analyze and assess the product, however, the difficulty is to understand what specific tools are used for what specific problem. He (E1) points out that is often what Philips struggles with.

Interviewee (E2), explains that the initiative towards structuring the NPD process includes the introduction of the Lean Startup method, in applicable situations. He describes the first stage (research) as liberal compared to the other stages, and researchers are encouraged to perform experiments. Additionally, interviewee (E2) describes the research stage as mostly exploratory, the next phase has as a goal to prove the principle, and to move even further it is required to prove the semi-product. Every innovation has to go through a funnel, some fast and some slower. Participant (E2) concludes by adding that the management team is supportive towards product champions, as long as they have the skills and competencies to be a part of the projects.

At Apollo Vredestein, participant (F) describes the product development process as a feedback loop which evaluates how the product is perceived and used based on acquired user feedback. Also, there is a formal Stage-Gate process for developing new products with typically a 2 year development cycle. Further, there are different groups in the process (concept group for generating new ideas, platform group that evaluates concepts, and product development group which is working out the details based on the platform) meaning that the end-products are derivatives of a platform which is built up of building blocks of concepts.

A6. Big Data in New Product Development

According to participant (A), Big Data is an important component since their product development is dependent on data. He (A) argues that companies need to begin at the end by first identifying the problem and evaluating if the problem is worth solving, otherwise there is a risk of committing resources, and creating technology for extraneous matters. By starting from the back - first, identifying and defining the problem - it is possible to forecast the value and usefulness of solving the specific problem, and subsequently test if valuable predictions can be made by using Big Data. Further, he (A) explains that organizations need to have a strategy on how to use the predictions, otherwise it is pointless to engage in such. Participant (A), explains that the data input which goes into to the software is constantly changing, so there is a steady need to tweak adjust their software. He (A) explains, as more data is made available and accessed - the company can perform R&D on larger datasets to develop the user-interface, evaluate and improve the software algorithm. Therefore, the product development and its functionality is dependent on Big Data.

Interviewee (B) states that Big Data can be used in every phase of the NPD process but each case or purpose would require a different approach. He argues that the utilization level of using Big Data in the NPD process, is directly affected by how data-driven the company is. Types of data further determine which technologies to use in order to cover the needs and achieve the specific goals. Finally, participant (B) argues that central data is more used in the formal product development phase rather than the rest.

According to participant (C), Thales is successful in utilizing and managing Big Data for its core markets. However, she explains that they face challenges in Big Data utilization for NPD directed towards new and emerging markets, where Thales is not yet successful.

According to interviewee (D1), Big Data has so far been more used for enhancing products and saving money than NPD. However, he (D1) argues when the data is part of the product you get different dynamics of using Big Data in R&D and product development. Product connectivity enables data feedback loops, which in turn improve product design, product development and maintenance. Moreover, interview participant (D1) at TNO claims that products without Big Data elements can still use Big Data to enable innovation processes.

Regarding the use of Big Data in the NPD process, participant (D2) argues that this fusion can improve efficiency and effectiveness for existing products. Furthermore, for using Big Data in the NPD process, it requires a process of discovering what the best way to achieve that implementation is. He (D2) adds that providing the R&D department with Big Data from multiple sources can lead to innovation discoveries. An important utilization arises from building-in feedback mechanisms for late-stage product development. (D2) concludes that Big Data is more valuable for technical applications rather than consumer applications, and believes that Big Data is great in theory but less impressive in real practice examples.

Participant (E1) explains that works with Big Data to identify a new need or new set of clients for the existing market or a totally new market. Further, Big Data is used to track activity and other elements within their product, such as items and infrastructure with internet connectivity and energy optimization. Gathering right data allows Philips to optimize products and development. Interviewee (E1) states that Big Data usage in the NPD is not being done on a large scale, and that the organization is still exploring the application of it.

According to interviewee (E2), data are utilized in research and FFE in order to generate new ideas and be introduced in the market. He explains that market data has been used for a long time, but connected data products that generate data and the process of exploiting that data is new. Big Data consists just one element of the factors that can affect the NPD, therefore the process will change a bit but in essence remain the same.

Participant (F) describes that analytics of large datasets are mostly utilized in the manufacturing or production phase. A lot of concepts will be in electronic parts and Big Data can set up the structure, especially if the concept is a data generating step. For that reason, Big Data is seen as a tool for analysis for many different projects and serve decision making in the product development process.

A7. Big Data Benefits

Interviewee (A), explains that a benefit from the Big Data trend is that top management is more informed of data potential. It is now seen as a resource to create value, and the top management pays attention to opportunities that have actually been existent for a long time. Data projects have now the support and interest from top management, which enables them to be carried out in the fear of missing out on value. He (A) regards data as the core feature to make predictions and improve the product in several occasions, which it would be impossible without their usage. Furthermore, he (A) adds that the usage of Big Data in the NPD process is a valuable tool to re-invent business models and exploit business opportunities.

Participant (B) acknowledges that the benefits of Big Data depend of the type of company and industry. However, general Big Data benefits in NPD are in the form of reduced time-to-market

(TTM), and being able to better adapt ideas, concepts and products on behavior to better fit with the demand. Interview participant (C) argues that the main benefit of Big Data is reflecting the creation of new markets so that Thales' can continuously develop products.

Enhanced market research and precise targeting of consumers are mentioned to be Big Data benefits by the interviewee (D1). Cost efficiency and targeting new partnerships are other benefits of using Big Data. Furthermore, Big Data can result in new products and improve the value chain with new insights. Finally, participant (D1) argues that Big Data allows for strategic positioning in the new emerging value eco-system.

Interviewee (D2) suggests that Big Data can significantly improve effectiveness and efficiency of producing and improving existing products. Further, participant (D2) suggested potential benefits from three different data approaches. The first approach, is the predictive approach which is helpful for new insights such as market trends. The second approach, involves prescriptive analytics and the ability to enhance the decision-making process, while the last approach is examining the validity of decisions and actions through descriptive analytics.

According to interviewee (E1), the company can enable cost reductions and efficiency with implementing Big Data. Further, right data allows for exploring new opportunities in their existing markets or entering new markets. The interview participant (E1) argues further that access and use of Big Data can boost not only internal factors but also make an impact and difference in society (smart city).

Participant (E2) explains that Big Data can enable the organization's insights in a digital form through user tests, and provide better understanding of what customers' feelings and needs are. Also, utilizing the potential of data can assist with the identification of what the organization should target. The value is in the strength of the ability to utilize it.

According to interviewee (F), Big Data has a lot of potential for products or development that can be easily tested in different circumstances. However, such potential is enabled only in the bigger context, if you can generate enough data.

A8. Managerial Challenges of using Big Data in the NPD process

One of the main concerns of participant (A) with regards to managerial challenges, underlies in the capability and perception of the organization to acquire necessary human resources, technologies or tools for Big Data. He (A) points out that the speed and volume of Big Data can be addressed by scaling up the technology, and being smarter on how to use the data and the analytic tools. Mistakes in Big Data projects are connected to the approach that companies will find value in the data, despite the fact that they do not understand the specific business case. Many projects will fail because organizations do not know how to use great data discoveries and turn them into monetary value. Another significant obstacle that many organizations face is the culture. He (A) explains that companies are organized around the human resources which possess various skills and expertise, but suddenly they deal with a lot of data that requires new skills and data understanding to receive their input. In order to get value from data, companies need to adjust their processes. It could be in the form of changing business models or changing the maintenance processes. Big Data experts need to be proactive and act like active consultants, due to the fact that management does not know how to implement the data and generate value from it. Managerial challenges of using Big Data also occur with regards to organizational cultures not being data-driven, and that the domain experts are too skeptical. Interviewee (A) further explains that every organization has unique problems, for the sake of his own company the challenges are in the form of cold start problems - not having enough data, since the product is dependent on the access to it.

Interviewee (B) points out that managerial challenges of using Big Data in the NPD are most often connected to the corporate culture and processes. Employees are reluctant to change and it translates into an obstacle since Big Data implementation is a long and on-going process. He (B) mentions that acquiring right expertise is a challenge. In the future, it will be an even bigger challenge due to the fact of the shortage of talent in the field. Interviewee (B) emphasizes the challenge of internal competition, where departments are not willing to share their data which in turn disables data combination and value-creation. Therefore, the managerial challenge lies in solving how to integrate data from silos and isolated departments into centralized systems.

The interviewee (C) specifies that challenges are now-a-days not only connected with Big Data but also Internet of Things (IoT) and the Cyberspace. She clarifies that challenges are becoming more complex, and being able to have a clear overview of decisions and impacts is more difficult. Such complexity further challenges decision-making processes. According to interviewee (C), for organizations to succeed, derive great value from Big Data and be upfront, they need to take more chances and expose themselves to risks. Furthermore, she argues that there are scarce human resources for Big Data objectives (e.g. data scientists). Another challenge connected to Big Data that is mentioned by interviewee (C) is concerning the usage of intangible assets (IPR) and Big Data.

According to interviewee (D1), scaling up new ideas is a challenge due to legacy systems. The changes that need to be made require organizational resources, capabilities and agility. Another major challenge for organizations to overcome is data ownership, interoperability issues, trust and privacy. Those need to be addressed to be able to share and connect data with internal and external actors.

Interviewee (D2), states that a main challenge of Big Data implementation in the NPD process, is what he calls the cultural clash, the inability of data experts and managers to understand the common goals. In addition, he believes that the top management should possess a basic knowledge of data technology as well as process and applications. Finally, an important enabler of Big Data implementation is the acquisition of the necessary resources, in the form of data scientists, data sources and technology.

Participant (E1) states that Big Data is a relatively new concept for Philips, he sees a lot of potential and thinks that more tools and techniques should be used than the analytic tools used in the past. Further, (E1) believes that it will be challenging to find the efficient methods for decision-making and to draw the correct conclusions. Finally, he urges the management team to take action in the matter of legislation, and compared the new rules that should be adopted with the legislation which allows Google to operate with Big Data.

Participant (E2) claims that getting the right data is the biggest challenge, since often the data is no yet there and you have to generate it. Furthermore, it is vital to have an interactive attitude and culture that the people are knowledgeable about the application and can interact well with these engineers and data scientists. Therefore, acquiring the right people as well as ensuring that they possess the right soft skills to go across their border and competencies field, is an important organizational challenge to address. Finally, interviewee (E2) suggests that the organization should create accessible and open information sources, in order to assist the work of human resources.

Interviewee (F) argues that an obstacle is to have sufficient amount of clean data which is rightly organized. Apollo Vredestein is still at a stage where they explore how to organize and evaluate the data, as well as investigate what can be done with it. There is a lot of potential of usage in plants but at the same time it is time and resource consuming to clean out the databases to measure more relevant data and do perform the right analysis in a systematic way. Another challenge according to participant (F) is the access of data from the eco-system of suppliers and users. From a top managerial perspective there needs to be an allocation of resources and a clear vision to implement the Big Data correctly into processes. Lastly, challenges that have to be addressed are cultural in nature. People need to work according to set procedures. However, the acceptance of people to use it and change already established procedures and tools is an obstacle, especially when you produce physical products with more conservative employees.

A9. Key Success Factors

According to interviewee (A), there are a lot of organizations that use bad technology and bad statistics. However, the technology that is being used is not the biggest success criteria. He (A) explains that key success factors for organizing for Big Data is: project management, experience using data in different projects, building and understanding the business case, as well as criteria that will makes any project successful are also key factors in Big Data projects. Interviewee (A) argues that data-driven culture, processes and project management is more important than the applied technology, since critical problems of using Big Data mostly occur within these domains. In addition, he (A) highlights that the organization needs to find a connection with business people and data people to be successful. Further, interviewee (A) specifies that the opportunity is in combining data, however, a success criteria is to develop a business case first, and then spot what the data is and test if you can use it for the specific case. Interviewee (A) summarizes that good quality data which is centralized and accessible, together with a developed business case within the organization and a culture that is data-driven are important success criteria of creating value from Big Data.

Participant (B) explains that one of the most important factors to succeed, is to have a centralized data source which will be easily accessed by all employees in the organization, hence enable them to combine the data. Further, he believes that this source should be expanded with more external data in order to increase insight and enable data experiments. Another important aspect of utilizing Big Data in the near future, is having specialized staff. Specifically, he (B) argues that there will be a shortage of human resources and that organizations should start training their current staff to be more Big Data oriented, since hiring the right expertise will be more costly.

Finally, interviewee (B) explains that that culture should be shaped in such a manner that data is involved in every process within the organization.

To succeed with implementation and organizing for Big Data in NPD, interviewee (C) states that a key success factor is to have multidisciplinary teams of specialists. Further, have engineers (which are generalists) that can oversee the entire scope, meaning that they oversee and understand operational, tactical, and political aspects of using Big Data for NPD. She argues that for companies to be successful they need to understand that handling of data has both direct and indirect influences. Lastly, she explains that the focus should not be on only handling data, but also to oversee all stakeholders and their interests in the data.

According to interviewee (D1) at TNO, having a clear Big Data objective and strategy is essential to successfully organize for Big Data. Further, having access to the data by cooperating with different stakeholders with shared ambitions and goals to enable sharing of data is a key element. Interviewee (D1) emphasizes the importance of speed and hyper-connectivity.

Participant (D2) considers the acceptance and adoption by the users, the people working with the outcomes of the Big Data analysis for the NPD, as an important factor of successfully implementing Big Data in the NPD process. He (D2), further believes that in order for Big Data technologies to be accepted, it is vital to construct the conditions and the right framing at the right time. Finally, it is important that people develop a mental model for accepting data innovation and analysis, and push towards privacy laws that allows Big Data innovation.

Interviewee (E1), points out the key factors for Philips to implement Big Data successfully. He (E1) argues that the staff needs to be properly trained and understand the data being offered as well as what technical and business actions they should perform to achieve the goals. Furthermore, it is crucial to analyze and draw conclusions as fast as possible and take advantage of the information extracted. As a final factor, (E1) refers to the need for a legislation change that will allow the proper use of Big Data which will benefit everyone, both Philips and the consumers.

Interviewee (E2) identified as an important factor the ability to obtain the right data and run sufficient data tests with them. Further, culture should promote openness to experiments and exploration of new things, but also promote a sense of urgency. Therefore, the importance underlies in having the right data science competencies in order to be able to build the first proof point early.

According to participant (F), a success factor is to prove a success story and have early adopters in the development team. There needs to take little time to do the proper analysis, which translates into the factor of having a good database with reliable connections to the data. Another important factor is to acquire sufficient resources, i.e. human- and technological resources. Lastly, participant (F) stretches the key factor of identifying if the Big Data approach is suitable for the problem the organization wants to solve.

Appendix B. Big Data enabling technologies and tools

Big Data in the Cloud	Most, if not all, of the Big Data technologies are completely connected with the cloud or operate with cloud versions. Stored Hadoop clusters are provided by several cloud services according to the customer need. Cloud technologies are enhancing Big Data with a significant value, since they support a rapid analysis in exchange for a relatively small cost. These two enablers, act as incentives for organizations to obtain more data and to seek for an augmented processing power and so on.
NoSQL Databases	Databases focusing on the storage and retrieval of high volumes of data. Their success underlies on bending the restrictions of conventional databases and result in scalability.
MapReduce	It is a programming paradigm based on two tasks, the “Map” task and the “Reduce” task. The first task is transformed into a different set value pairs or tuples, while the second task combines output of the “Map” task to create a reduced set of tuples. It enables job execution scalability against a large amount of servers.
Storage Technologies	The rapid technological advancement in the volume of data, requires efficient storage capabilities. The progress in this sector is mostly focused on data compression and storage virtualization.
Column-oriented databases	They store data on columns rather than rows, resulting to a rapid process and data compression. Limited by batch updates and a time-consuming update.
Hive	Originally developed by Facebook, Hive acts as an “SQL-like” link that enables Business Intelligence applications to run queries against a Hadoop cluster as if they were manipulating a traditional data store, and also enhance the effectiveness of Hadoop.
PIG	Acts as a link between for Hadoop to connect developer and business user realities. The difference from Hive underlies to the fact that PIG is not using an “SQL-like” language and enables query execution over Hadoop cluster data. Open-source, originally developed by Yahoo!.
WibiData	WibiData is a merger of web analytics and Hadoop, built on top of HBase, and itself a database surface on top of Hadoop. It focuses on real-time applications, personalization and recommendations to users.
PLATFORA	Hadoop’s biggest limitation due to the low-level implementation of MapReduce. It demands a deep developer knowledge to be managed, it is time-consuming, and it transforms queries into Hadoop jobs.
SkyTree	SkyTree is a Big Data-focused platform, characterized by a high effectiveness on machine learning and data analytics. Machine learning has a significant role, since manual processing of the large amount of data is either

	too expensive or impossible.

Appendix C. Interview Guide

Company & Business Background

- What is your position within the company?
- How is your role related to Big Data and/or Business Development? [Business Intelligence, Business Analytics, Innovation Management]

Big Data Understanding

- How would you define Big Data?
- How does your company engage in Big Data: capture and collect data?
- How does your company manage the data and verify its quality?
- What is your company's Big Data/ data mining objective?
- What types of data is considered useful?
- How do you select and assess the potential value of data?
- What are the success criteria for using Big Data?
- What is the organizational setup for managing of Big Data? [Unit Structure, Responsibility]

New Product Development

- Can you describe your company's process for New Product Development?
- How do you structure/manage the Pre-development/Front End of the Process, e.g.
    - Idea generation - idea selection - opportunity analysis - opportunity analysis - concept definition - concept selection
- What tools do you use to evaluate and analyze your projects throughout the development?

Big Data & New Product Development

- How does your company make use/implement Big Data in the NPD process?
    - or, what function does Big Data have throughout the NPD process?
- In which phase is Big Data used? [map the Big Data usage across the entire NPD process with pre-development and launch]
- What Benefits/potential have you noticed of using Big Data in your NPD process?
- How would you rate your company's utilization of Big Data in New Product Development?
    - Successful: Can you quantify some of the results?
    - Unsuccessful: Can you identify the reasons?
- Do you have/had any actions planned to improve the management support towards Big Data implementation in the NPD process?
- What managerial challenges have occurred/exist connected with the usage of Big Data in the NPD process?

    - Top management’s involvement
    - Cross-functional collaboration
    - Resources and Capabilities
- Are there any key success factors of using Big Data in the NPD process?
    - Organizational
    - Structural

Recommendations & Future Research

- Can you provide us with internal or external contacts for the purposes of this research?
- Do you wish to add anything else to our research?