MASTER THESIS

CBS COPENHAGEN BUSINESS SCHOOL HANDELSHØJSKOLEN

Big Data for Big Ideas: Exploring the Potential of Big Data Analytics for New Product Development

Authors:
Sam Anteney-Hoare (677476)
Maximilian Krempel (676885)

Supervisor:
Vera Rocha

Pages: 119
Characters: 271,008

Date of Submission:
11th September 2019

M.Sc. in Management of Innovation and Business Development (cand.merc.)

[Page intentionally left blank]

Abstract

New Product Development is facing pressures to innovate faster and at lower cost whilst ensuring a better fit to the market's needs. At the same time, companies are becoming increasingly immersed in Big Data, the analysis of which represents a potential pathway to overcome these pressures. However, the research area covering the intersection between Big Data Analytics and New Product Development is still in its infancy and fails to provide a holistic perspective which can inform practitioners and researchers alike. This thesis aims to provide such a perspective and explores the potential of Big Data Analytics for New Product Development. In particular, it investigates how Big Data Analytics can influence New Product Development performance and identifies how this influence is contingent on organisational and contextual factors. In order to fulfil the research aim, this thesis followed an inductive approach, applying the Grounded Theory Method to analyse 12 in-depth interviews with experts from consultancies, product companies and academia. This thesis finds that Big Data Analytics has the potential to positively influence New Product Development performance and provides a typology consisting of 11 themes which showcase the means by which this can be achieved. Further, the positive influence on New Product Development performance is greater with higher levels of organisational agility and with a lower degree of innovativeness. Finally, the industry context was included *ex-post* due to its contingent role in the explored relationship. Through its findings and typology, this thesis extends the current patchwork of studies, providing structure to this emerging research area and offering practitioners a state-of-the-art overview on the intersection of Big Data Analytics and New Product Development.

Table of Contents

Table of Figures

Table of Tables

Glossary & Abbreviations

AI	Artificial Intelligence
BD	Big Data
BDA	Big Data Analytics
BDAC	Big Data Analytics Capability
GTM	Grounded Theory Method
IoT	Internet of Things
MDAC	Marketing-enabled Data Analytics Capability
NPD	New Product Development

1 Introduction

In his seminal book *“Capitalism, socialism and democracy”*, Schumpeter (1942, p. 83) argued that a process of *“creative destruction is the essential fact about capitalism”*. Underpinning this theory is the notion that all companies must adapt and innovate to meet changing environments and to stay ahead of the competition (D’Aveni et al., 2010). As firms compete in ever more volatile environments (Fagerberg et al., 2004; Johnson et al., 2017), authors acknowledge that the process of innovation is *“the core renewal process in any organization”* without which a firm risks its survival and growth prospects (Bessant et al., 2005, p. 1366). The view that innovation is crucial for a firm’s success extends to product innovation which is cited as a decisive factor in the sustainable success of a company's business activities (Henard & Szymanski, 2003). To exemplify this, products that have been developed within the last five years represent almost one third (or more) of the amount of sales and profits of many companies across a wide range of industries (Schilling, 2013). Furthermore, the importance of product innovation has transcended the academic community and is also widely recognised by practitioners. In the annual Global Innovation Survey, conducted by the Boston Consulting Group, 79% of CEOs consider product innovation to be one of their top three strategic priorities (Ringel et al., 2015), an unprecedented figure since the study was first conducted in 2005.

While increasingly volatile environments are further driving the need to continuously innovate, the process to generate product innovations, known as New Product Development (NPD), simultaneously faces severe challenges. First, the phenomenon of globalisation is ensuring that markets become progressively more interconnected generating increased information flows. The increased information flows around the globe lead to more well-informed customers adopting new technologies more frequently (Palacios Fenech & Tellis, 2016), and at the same time, faster imitation of new product innovations by competitors (Canuto, 2018; Stremersch et al., 2010). For companies, the consequence of this is less time to profit from a new product innovation, before it is either imitated or a newer version released (Palacios Fenech & Tellis, 2016). Effectively this means that product life cycles are being shortened (J. Chen et al., 2012). Estimates suggest that over the past 30 years product life cycles have been cut to one third or even one quarter (Trinkfass, 1997), while development times are still perceived to be too long (Ringel et al., 2015). This shift requires companies to continuously innovate at a faster pace (Johnson et al., 2017; Liu & Kop, 2016).

Second, more intense global competition, resulting from globalisation, is dragging many product companies into a situation whereby they are unable to differentiate themselves and so suffer downward pressure on their prices and profit margins. This situation is referred to as the commodity trap (Chesbrough, 2019). At the same time, customers want new products at unprecedented low prices and show little brand loyalty, constantly comparing products across different companies (Cooper, 2014). The consequence of this for NPD is that there is an increased focus on controlling the costs of development for new product launches (Zhan et al., 2017). Put differently, if a firm is forced to react to the outlined pressure and reduce its prices for new products then in order to maintain a similar return on investment from the innovation process the costs must be similarly reduced. However, the evidence suggests that this is very difficult to achieve as companies today are *"devoting approximately 20 times as many people to R&D as their peers did in 1930, but the output from all of this endeavour is not rising in unison"* (Gaskell, 2018).

Third, whilst companies attempt to innovate faster and halt the rise in product development costs, they are faced with the reality that many new products launched into the market fail as their attributes do not meet customer requirements (Schilling & Hill, 1998). Estimates vary widely with the popular press reporting figures between 80-90% for product failure rates, whilst empirical studies suggest that the figure is closer to 45% (Castellion & Markham, 2013). However, even given the 45% failure rate this means that there is high pressure on those products that do succeed to recoup the investments in failed products. Surprisingly, given the increased focus on identifying NPD success factors in academia, the failure rate of new products has not improved since the 1980s (Castellion & Markham, 2013).

Taken together, the pressures resulting from globalisation are creating a need to innovate faster, to meet shorter product life cycles, and at lower cost, to maintain sufficient innovation returns, whilst ensuring a better fit to the market's needs, to reduce the threat of new product failure. In short, firms must overcome these key challenges to improve their NPD performance, which is fundamental to their continued survival and growth. Therefore, attention should be paid immediately to how organisations can innovate products more effectively (Zhan et al., 2017).

At the same time, the emergence of new technologies and platforms, such as the Internet of Things (IoT), social media or multimedia, and the continued trend towards digitalisation has led to increasing amounts of data in all aspects of society. In fact, it is estimated that today 98% of the world's information is digital and thereby represents some form of data (Xu et al., 2016) and that this amount will continue to increase exponentially (Desjardins, 2019).

The abundance of data has led scholars to suggest that many companies now operate in "*data-rich environments"* (N. Bharadwaj & Noble, 2017, p. 476). There are many clear benefits to the use of Big Data (BD) in business and some scholars go so far as to say that it could be described as a *"management revolution"* as it enables managers to make decisions based on evidence rather than their own intuition (McAfee & Brynjolfsson, 2012, p. 62). Whereas these benefits of BD have in the past been used to optimise operational processes, increasingly authors are now acknowledging its potential to be used in the creation of innovations (Davenport, 2013; Gobble, 2013; Kiron et al., 2012). The transition to a data-rich environment is seen not only as an important factor for innovation in general, but in particular also for the successful development of new products (N. Bharadwaj & Noble, 2017). With respect to product innovations, the use of BD and the analysis of such data, known as Big Data Analytics (BDA), can improve the performance of crucial activities in the NPD process, such as deriving information about customers or testing prototypes (Xu et al., 2016) and can enable "*companies to come up with genuinely innovative new products"* (Zhan et al., 2018, p. 592).

Despite the increasing interest, however, researchers believe that most of BDA's potential for NPD is yet to be discovered (Zhan et al., 2018). Hence, scholars have highlighted the need to better understand the impact of using BDA on newly developed products (Citrin et al., 2007). The limited literature in this field can be classified into two areas. On the one hand, a few studies have investigated the use of BDA for some specific activities in the NPD process, such as idea spotting in online communities (Christensen et al., 2017) or the identification of preferences through social media (Lee & Bradlow, 2011). However, these studies mainly focus on the Ideation stage and are also driven by academic research rather than representing concrete use cases that are already leveraged in practice. Against this backdrop, Mikalef et al. (2019) outline that while there is growing literature on the general business potential of BDA, there is still a lack of empirical research on the mechanisms and conditions under which BDA yield the biggest potential for creating business value. The result of this is that practitioners are left in *"unchartered territories when faced with implementing such investments in their firms"* and so it is not surprising that only a few companies have been able to capture the full potential of their BDA investments (Mikalef et al., 2018, p. 548). Consequently, the means by which companies can incorporate BDA into their NPD process are still unclear, specifically in which activities and stages of the NPD process BDA can be helpful.

On the other hand, studies have investigated the relationship between BD and NPD, identifying success factors for improving NPD (Zhan et al., 2017), or have explored the relationship between BD and NPD related themes, such as exploration and exploitation activities (Johnson et al., 2017). However, to the best of the authors' knowledge, so far there has been no study investigating if and how incorporating BDA can influence NPD performance and thereby can represent a way to react to the outlined pressure to innovate more effectively. With the intention to derive an empirically grounded typology for incorporating BDA in the NPD process and to analyse the effect of incorporating BDA on NPD performance, the first research question is:

RQ1: How can big data analytics (BDA) influence new product development (NPD) performance?

1a: What are the means by which BDA can be incorporated into NPD?

1b: How can incorporating BDA influence NPD performance criteria?

Whilst incorporating BDA may lead to improved insights for the NPD process, it is by no means certain that these insights will lead to the creation of business value. As Mikalef et al. (2019, p. 293) state, *"data-driven insight is only one component of gaining value from big data investments, the other is responsiveness"*. In other words, responsiveness is referring to a firm's ability to respond to the changes highlighted by the data-driven insights. Davenport et al. (2012) echo this sentiment and argue that against the backdrop of fast changing environments highlighted above, it becomes even more important for companies to react quickly and intelligently based on the identified insights. Literature within the information systems field refers to 'responsiveness' as *organisational agility* (Sambamurthy et al., 2003; Shuradze et al., 2018; Tallon & Pinsonneault, 2011) however there are very few studies which examine the influence that organisational agility has on the use of BDA in the innovation process. Shuradze et al. (2018) find that organisational agility is a mechanism by which companies can move from data-driven insights to effective innovation actions. While their study aimed to empirically establish the link between organisational agility and innovation success, there is little guidance as to how this relationship functions in reality. As a result, it is still hard to ascertain why organisational agility is so important for innovation success and what specific elements of innovation success are affected, and thus it continues to be a relationship that needs to be looked at in more detail.

Moreover, the NPD process is not always standardised, which in turn could have an influence on how BDA is incorporated and what value can be derived from its use. The degree of innovativeness of the product innovation has been shown to alter the NPD process, with radical innovations making it longer and messier than for incremental innovations (Veryzer Jr., 1998). Despite these differences, Mikalef et al. (2019) find that BDA has a positive effect on both, incremental and radical innovation capabilities. Whilst their study empirically proves a link between BDA and product innovativeness there is still a lack of research as to *how* product innovativeness affects the use and value of BDA for NPD (N. Bharadwaj, 2018).

In summary, the extent to which BDA influences NPD performance appears to be contingent on certain organisational factors. Therefore, based on these considerations, the second research question that this thesis addresses is:

> **RQ2:** How is the relationship between BDA and NPD performance contingent on organisational factors?
>
> > 2a: How is the relationship between BDA and NPD performance contingent on the level of organisational agility?
> >
> > 2b: How is the relationship between BDA and NPD performance contingent on the degree of innovativeness of the product innovation?

In order to answer the two research questions outlined above this thesis will follow an inductive approach, specifically, the Grounded Theory Method (GTM). This method was deemed most suitable for this thesis given that the area of investigation is relatively understudied and consequently the focus is on developing new theory which is anchored in the collected data (Glaser & Strauss, 2006). The analysis is based upon qualitative data, which was collected via two rounds of 12 interviews with representatives from leading consultancies, product companies and academia. Following the GTM each interview builds upon the one before allowing for the research scope to be narrowed and trends in the data to appear (Glaser & Strauss, 2006).

The contribution of this thesis can be described as manifold. First, it aims to develop a typology based on real use cases showing the different means by which BDA can be incorporated into the NPD process, thereby providing structure to this currently understudied and fragmented research area. Second, this thesis aims to explore the influence that BDA can have on three NPD performance criteria, namely cost of development, speed-to-market and product-market fit and thus elucidate the potential that BDA can have for NPD. Third, this thesis aims to examine whether and how the relationship between BDA and NPD performance is contingent on the two organisational factors of organisational agility and the degree of innovativeness.

The remainder of the thesis will be organised as follows. After the introduction, chapter 2 will outline the methodology employed to guide this study. Chapter 3 will explain the appropriate literature related to BDA and NPD separately as well as the currently limited literature on the intersection of these two distinct research areas. In addition, the literature on potential contingency factors is also described here. Chapter 4 will present the conceptual framework underlying this thesis, which was developed based on literature and initial data collection. In this chapter the constructs under investigation are also defined and operationalised. Chapter 5 will present the main results and analysis of the thesis covering both main research questions. Chapter 6 will discuss and critically examine the key findings of the analysis and will indicate practical and theoretical implications as well as limitations and directions for further research. Chapter 7 will conclude this thesis and summarise the central findings.

2 Methodology

In line with Guba and Lincoln (1994) and Saunders et al. (2009), this thesis first addresses the research philosophy, and thereby the questions of ontology, epistemology and axiology as they represent the fundamental paradigms through which research is perceived. Subsequently, the research approach and design of this work is presented. Thereafter, the method applied in this master thesis is stated and described in detail. Finally, the methodological limitations in terms of reliability, generalisability and validity are outlined.

2.1 Research Philosophy

Ontology deals with the researcher's view of the nature of reality (Saunders et al., 2009). In this context, two different views can be distinguished: objectivism and subjectivism. Objectivism affirms that social phenomena have an existence that is independent of, and external to, social actors (Bryman & Bell, 2011). Through an objectivism lens, an organization is seen as a tangible object with a reality that is external to the individuals that are related to it and, as a result, the organization actually guides and constrains the actions of its social actors (Bryman & Bell, 2011). In contrast, subjectivism portrays the position that the existence of social entities derives from the perceptions and subsequent actions of social actors (Saunders et al., 2009). The subjectivism perspective would suggest that an organization only exists as a result of the actions of its social actors and consequently it is continuously evolving (Bryman & Bell, 2011). Within the context of this master thesis, the authors do not want to commit themselves strictly to one of the two perspectives, but rather interpret research philosophy in line with Tashakkori and Teddlie (1998) as a continuum instead of two mutually exclusive positions. On this continuum the authors lean towards the objectivistic perspective, but due to the ambiguity in defining a clear position, the authors' perspective on ontology could be described as that of pragmatism (Saunders et al., 2009)

Epistemology is concerned with what is (or should be) regarded as valid knowledge in a field of study (Bryman & Bell, 2011; Saunders et al., 2009). A central issue is whether the social sciences should be studied according to the same principles and procedures as the natural sciences (Bryman & Bell, 2011). In this context, two research paradigms, known as positivism and interpretivism, take contrasting views on social sciences (Collis & Hussey, 2003; Saunders et al., 2009). Positivism derives from natural science and claims that social settings can be analysed objectively. However, the exclusion of social circumstances eventually led to the development of the interpretivism research philosophy, which endorses social complexities. Interpretivism essentially believes that the subject matter of social sciences (i.e. people and their various institutions) is fundamentally different from the natural sciences and consequently requires an altogether different research logic (Bryman & Bell, 2011). Bryman and Bell (2011, p. 16) succinctly summarize the contrast between the two research paradigms as essentially a *“division between an emphasis on the explanation of human behaviour that is the chief ingredient of the positivist approach […] and the understanding of human behaviour. The latter is concerned with the empathic understanding of human action rather than with the forces that are deemed to act on it”*. This thesis adopts an interpretivist approach as the authors acknowledge that the findings are both subjective and influenced by individual perceptions (Collis & Hussey, 2003). However, this interpretivist approach also allows the authors to study the field of research in detail (Saunders et al., 2009).

Axiology describes researchers’ view on the role of values in research (Saunders et al., 2009). According to Heron (1996), values are the guiding reason for all human action and will thus also affect one's research, from the selection of the research question to data collection and the subsequent analysis. Heron (1996) proposes that researchers write a statement of personal values with regards to the studied topic. At this point, such a statement would go beyond the scope of this work, but the authors would like to give a few examples of the way in which personal values have influenced this research. With respect to the selection of the research topic, the authors’ curiosity for new and understudied research areas, as well as their enthusiasm for the fields of innovation and technology are decisive. The authors prefer to study topics in depth and explore the underlying rationale behind actions and decisions, rather than just establishing a relationship between different variables, and as such, personal interviews dominate the data collection process. In conclusion, it can therefore be said that the authors acknowledge that research is rather value laden and biased by the authors’ world views and experiences (Saunders et al., 2009).

2.2 Research Approach

There are generally two types of research approaches, namely a deductive and an inductive approach, which differ in the way they apply theory and observation (Saunders et al., 2009). A deductive approach takes advantage of existing theories at the beginning and tests whether they hold in practice. Conversely, an inductive approach begins with observations and seeks to develop these observations into a theory (Saunders et al., 2009). This thesis shall follow an inductive approach as the chosen research area is understudied and therefore there is little theory to base the underlying research questions on. Instead of testing if an existing theory holds in practice, this master thesis aims to make a theoretical contribution by developing a typology to help structure this emerging research field.

2.3 Research Design

According to Bryman and Bell (2011) the research design is a conceptual structure that will support applied research methods and the subsequent analysis of the generated data. Decisive in the election of a certain conceptual structure is the purpose of the research, which can be classified as either explanatory, descriptive and exploratory studies (Saunders et al., 2009). For studies that want to explain a new phenomenon and derive novel insights, the exploratory study is the appropriate mean (Robson, 2002). As this master thesis seeks to explore a phenomenon that can already be found in practice but is barely explained in the literature, it will be conceptualized as an exploratory study. This exploratory character is also decisive when considering using a qualitative or quantitative research method. Quantitative research describes data collection techniques and data analysis procedures that focus on numerical data (Saunders et al., 2009). In contrast, qualitative research collects and analyses non-numerical data (Saunders et al., 2009). In a setting where a research area needs to be explored and variables need to be discovered, rather than tested, a qualitative approach is better suited (Corbin & Strauss, 1990; Willig, 2001) and therefore will be the approach applied in this master thesis.

From the various qualitative research methods, the GTM was chosen against the background of the aforementioned research philosophy, approach and purpose. The GTM is used in research to generate theory using qualitative data. Consequently, this method is not suited to confirm existing theoretical concepts but aims to establish new theory that is anchored in the collected and examined data (Glaser & Strauss, 2006). In the GTM, data collection begins without a first theoretical framework having been created (Saunders et al., 2009). The method is characterised by a temporal parallelism and functional dependence of data collection, data analysis and theory building (Glaser & Strauss, 2006). Through this functional dependency, information that has already been collected and analysed can be verified by means of further data. In turn, the further procedure and the deepening data collection are based on the past results (Glaser & Strauss, 2006). In this iterative process, theoretical sampling, which means that the subjects of the further investigation must be selected in such a way that they advance the formation of theory, is an important component (Glaser & Strauss, 2006).

2.4 Research Method

2.4.1 Data Collection

To address the research questions of this thesis, the authors conducted a total of 12 interviews. Due to the size, the sample can be described as a nonprobability sample. This decision was made deliberately because this study does not attempt to make generalizable statements based on statistical calculations (Collis & Hussey, 2003). Further, smaller sample sizes can be considered very suitable for research using an inductive approach, such as this study (Saunders et al., 2009). When possible, interviews were conducted in-person to better record the non-verbal communication and to enable a more thorough investigation around areas of uncertainty, hesitation or contradiction. Ten of the interviews were conducted in person and two were conducted via Skype.

Due to the interplay between data collection and theory building according to GTM, data was not only gathered at one point in time but continuously between May and August 2019. In line with the concept of theoretical sampling, the choice of further interview partners was constantly refined and adjusted based on the previous interviews and the evolving theory (Glaser & Strauss, 2006). Despite the continuous nature of the data collection process of this thesis, it can be generally divided into two distinct phases. Table 1 gives an overview of the conducted interviews broken down by data collection phases.

Table 1: Breakdown of interview partners by data collection phase

Name of interviewee	Position	Company	Type of Organisation (Industry)	Geography	Interview Date	Duration
1st phase of data collection						
Kiran Vas	Co-Founder and Partner	2021.ai	Consultancy	Denmark	03/05/2019	54:47
Adam Hede	Management Consultant in Data with Impact Team	Implement Consulting Group	Consultancy	Denmark	13/05/2019	58:30
Abayomi Baiyere	Assistant Professor	Copenhagen Business School	Academia	Denmark	14/05/2019	25:15
Mathias Blom	Senior Data Scientist	Boston Consulting Group (Gamma)	Consultancy	Denmark	21/05/2019	53:20
Asbjørn Simonen-Andersen	Management Consultant in Data with Impact Team	Implement Consulting Group	Consultancy	Denmark	29/05/2019	53:49
2nd phase of data collection						
Rahul Shah	Partner	QVARTZ Analytics	Consultancy	Denmark	11/07/2019	1:05:08
Isadora Anjos	Senior Pipeline and Product Manager	Coloplast	Product Company (MedTech)	Denmark	15/07/2019	59:57
Patrick Ahlbrand	Corporate Product Strategy	CLAAS	Product Company (Machinery)	Germany	16/07/2019	37:55
Hannah Haugbølle Thomsen	Head of Development and Analysis in Product Management	Pandora	Product Company (Fashion)	Denmark	23/07/2019	1:03:07
Sebastian Barfort	Management Consultant	ReD Associates	Consultancy	Denmark	26/07/2019	59:47
Nicolas Antille	Statistics for Quality and Product Network Leader/ Virtual Prototyping	Nestlé	Product Company (FMCG)	Switzerland	13/08/2019	35:49
Dr. Nils Dülfer	Manager	IMP³rove Academy: European Innovation Management Academy	Consultancy	Germany	16/08/2019	1:06:48

The *first phase* was of an explorative nature and the aim was to broadly explore the link between BDA and innovation. This should not be seen as an absence of direction, rather it is essential for explorative studies as the initially broad research focus will be further narrowed throughout the course of the research (Adams & Schvaneveldt, 1991). The interview partners for the first phase were selected according to this intention. The majority of the interview partners (four out of five) were working in leading consultancies. The authors' reasoning behind this decision was twofold: first, consultants are experts in their field and are required to be aware of the state of the art by virtue of their role as consultants. Second, consultants are working with diverse companies of different sizes and from different industries and thereby were perfectly suited to provide an overarching perspective on our research area. Three out of the four interviewees from consultancies are experts in data analytics. The fourth interview partner is an expert in innovation management. These four interviews were complemented by an interview with an academic researcher from one of Europe's leading universities who specialized in the intersection between innovation management and digitization. This interview provided a cutting-edge, theoretical view on this understudied research area, as well as giving an indication about which individual aspects of the interaction between BDA and innovation could be worth further investigation. The interviews during this first phase were conducted in an unstructured manner. Due to the explorative nature of these five interviews, the authors refrained from directing the discussion in favour of openly exploring the topics together with the interview partners. See Appendix 1 for an example interview guide used for the first phase of data collection.

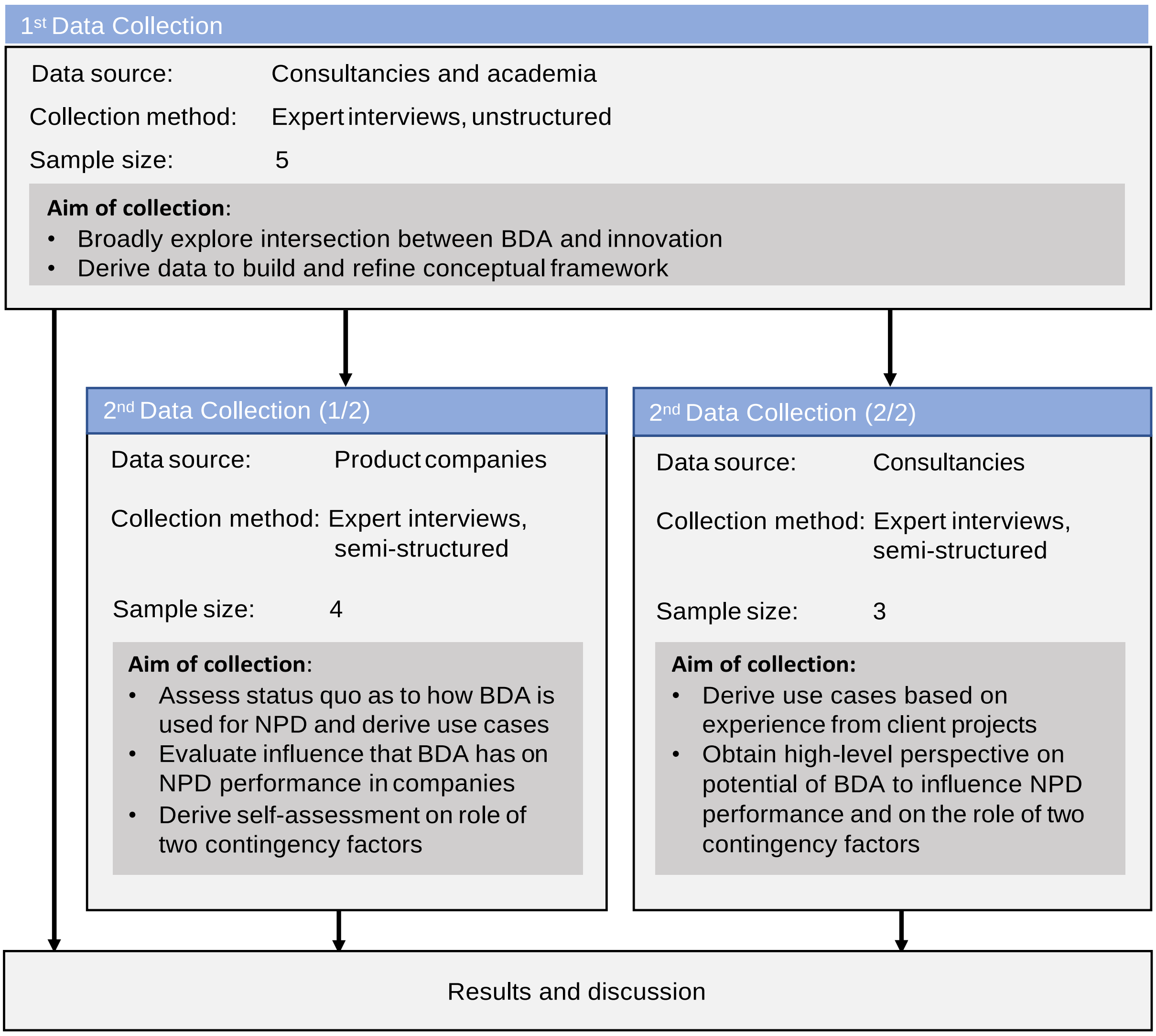
After the first phase of data collection, the collected data was analysed and complemented by a comprehensive literature review. Based on both empirical evidence and literature, the research focus was narrowed down to investigate the relationship between BDA and NPD performance and how this relationship is influenced by two organisational factors. Out of the many potential contingency factors, it was possible to limit the selection to the ones that seemed most relevant, namely the level of organisational agility and the degree of innovativeness of the product innovation. Accordingly, a conceptual framework was developed, which will be explained in detail in chapter 4.

In line with Miles and Huberman (1994) the authors regard a conceptual framework as a way to graphically, and in narrative form, present their understanding of the phenomenon and central constructs under investigation. This conceptual framework then formed the basis for the *second phase* of data collection, including the selection of the interview partners as well as the creation of a second interview guide. This second phase of data collection was targeted towards examining the different research questions as outlined in the conceptual framework. For this purpose, interview partners from two different areas were chosen.

On the one hand, interviews were conducted with representatives from product companies, giving in-depth insights on the status quo of how BDA can be used in the NPD process in practice. In order to identify suitable interview partners for this sub-sample, a multistage approach was followed. In the first step, companies were identified according to the criteria of being both a product company and also being data-driven, based on secondary information. In a second step, these companies were contacted, and a proposal was sent out describing the scope of the authors' research. Only if companies agreed that they already use BDA for NPD and hence represent a good fit for the outlined research topic, would the authors look to schedule an interview with these companies. When selecting the interview partners, the authors ensured that the interviewee worked in product development (or a similar department) and had some experience with data analytics themselves. The resulting sample of interviewees represents companies across a range of different industries, such as machinery, MedTech, fashion or FMCG. On the other hand, consultants specialized on the fields of innovation and data analytics, were selected to give in-depth insights into specific elements of the conceptual framework and to validate the information provided by the product company representatives without having a particular industry bias.

This second data collection phase followed a semi-structured interview approach and a specific interview guide was produced for each interview (see Appendices 2 and 3). The interview guides generally consisted of two overarching sections as well as introductory questions to establish the context of the interview and create rapport between the interviewee and the interviewers. In the first section the interview partners were asked questions relating to the first research question, specifically how, in their view, BDA can be incorporated into the NPD process and how incorporating BDA could influence NPD performance. As part of this, interview partners from product companies were asked for specific use cases on how BDA is already be incorporated into the companies NPD process and what influence that has on NPD performance. In turn, consultants were asked to provide use cases from client projects and to allude to the potential influence of incorporating BDA on NPD performance from a high-level perspective. The second section covered the second research question and focused on the two contingency factors, organisational agility and the degree of innovativeness, which could influence the relationship between BDA and NPD performance. In this context, product companies were asked for a self-assessment and thereby in how far the two contingency factors influence the relationship between BDA and NPD performance in their individual case, while consultants were again asked for a more holistic perspective based on their client experience. In general, the interview guide acted as a rough outline for the interview and specific follow-up questions, or floating prompts (McCracken, 1988), were asked to discover implications of answers provided and to clarify ambiguities. However, in general the interview guide was more strictly followed than in the first phase, due to the narrower focus of the second phase of data collection. The following illustration summarizes the data collection process.

Figure 1: Overview of data collection process (own illustration)

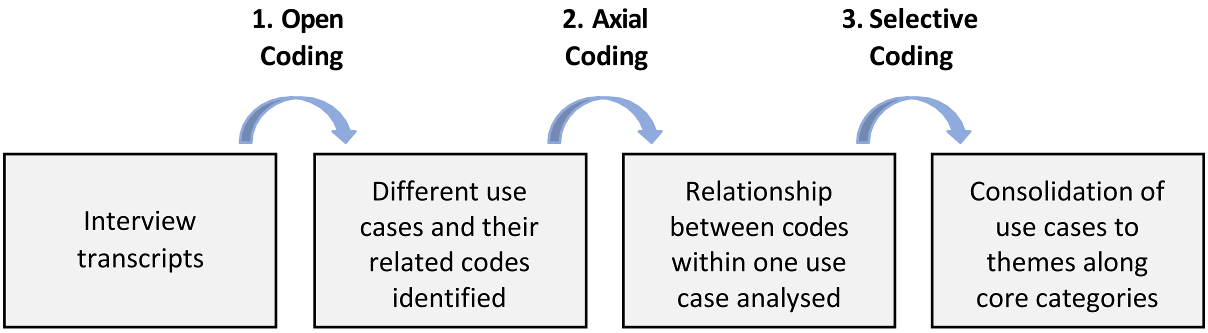


Throughout the data collection phases, all interviews were recorded and transcribed with permission (Yin, 2009). Transcription was conducted as soon as feasibly possible after each interview was held so that a more accurate depiction of the interview could be captured. Once transcribed each interview was checked independently to ensure that the transcription reflected the reality of the interview as much as possible. Upon completion of this process the extended text was then inputted into the NVivo qualitative research software for coding.

2.4.2 Coding & Data Analysis Process

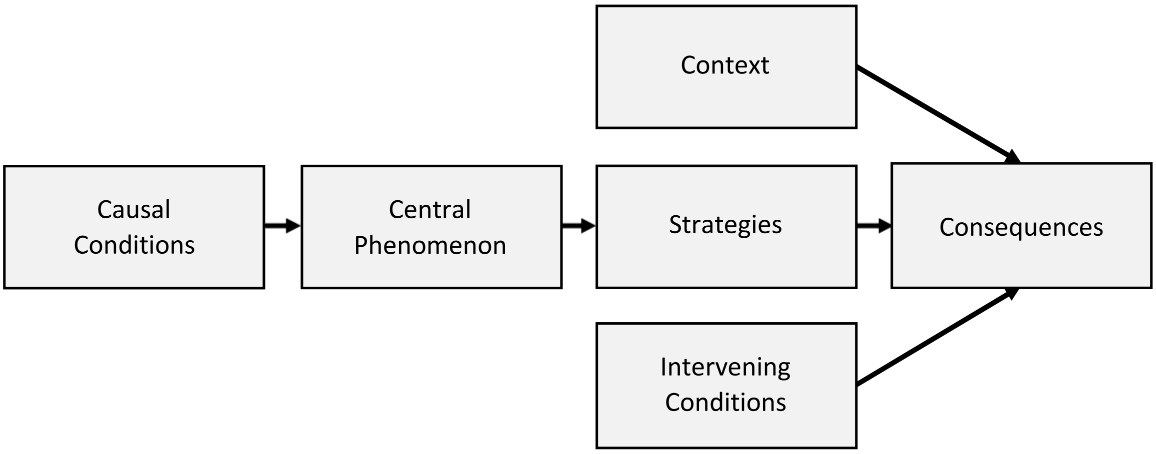
According to Charmaz (2006, p. 43), coding means *"categorizing segments of data with a short name that simultaneously summarizes and accounts for each piece of data"*. In the context of GTM, coding represents the analytic frame on which the authors' analysis will be based. Against the backdrop of the explorative nature of this thesis, coding plays an important role, as it not only sheds light on the predefined research area but can also lead to unforeseen research areas and even open up new research questions (Charmaz, 2006). As proposed by Corbin and Strauss (1990), open, axial and selective coding was conducted, even though the latter two were only performed for the development of the typology. Figure 2 provides an overview of the coding process followed for the development of the typology.

Figure 2: Three-step coding process for typology (own illustration)



Open coding represents the first step of the analysis in GTM and describes the process in which the data is initially broken down analytically. This type of coding was conducted with the complete interview transcripts of all 12 interviews. For this thesis, incident-to-incident coding was followed due to the large amount of transcribed interview material and the limited time horizon of this work. While coding, the authors considered the remarks of Charmaz (2006) – to keep codes simple, precise and short and to move quickly through the data. As part of the open coding, use cases in which BDA has been successfully incorporated into the NPD process were identified in the empirical data. Further, relevant statements about different performance criteria were identified and grouped according to the influence that incorporating BDA has on these criteria (i.e. positive or negative). Similarly, relevant statements about the two contingency factors were identified and grouped according to the indication given by the interview partners as to the influence that these factors have on the main relationship between BDA and NPD performance. After the open coding, axial coding was conducted with the use cases and related codes with the aim to progress with the development of the typology. In this second step, codes related to one use case were evaluated and brought into a relationship with each other (Corbin & Strauss, 1990). For that, the authors relied on a coding paradigm (see Fig. 3), which depicts the interrelationship between causal and intervening conditions, context, strategies and consequences in relation to one central phenomenon, in our case one specific use case (Creswell, 2012).

Figure 3: Axial coding paradigm (own illustration based on Creswell, 2012)



After the axial coding, selective coding was conducted in which theoretical concepts, hereafter referred to as themes, are built based on core categories (Corbin & Strauss, 1990). Different use cases were then combined and grouped into themes based on possessing similar core categories. The themes represent the different means by which BDA can be incorporated into NPD and can thereby influence NPD performance. Appendix 4 contains the output from the NVivo software after the selective coding, which depicts the coded use cases, grouped into the 11 themes and arranged by NPD stage. In line with the GTM, the authors generally maintained an open approach to data analysis and to the inclusion of possible emerging issues in the work. As a result, a further contingency factor could be identified in the course of the analysis.

2.5 Assessment of Methodological Limitations

According to Saunders et al. (2009) it is important to assess the credibility of the chosen research design. Therefore, the *validity*, *reliability* and *generalisability* of the research design have to be analysed.

Reliability describes to what extent the data collection techniques and analysis procedures can lead to consistent findings (Saunders et al., 2009). As the interviews were conducted in either an unstructured or semi-structured way, they naturally evolved around what was said, rather than strictly following an interview guide. Furthermore, the interviews were almost exclusively conducted jointly by both authors and so the possibility of observer error has to be factored in as the authors would naturally elicit responses in different ways (Saunders et al., 2009). Therefore, the reliability of the data collection can be assessed as being rather low. However, these more unstructured interviews allowed for a richer data set to be collected and are especially well suited for exploratory studies (Saunders et al., 2009). Hence the lower reliability cannot be seen as a weakness of the research design. In addition, the replicability of research in dynamic and rapidly evolving environments such as the one under observation is limited regardless, as the data always reflects the reality at the time of the survey, which is subject to change (Marshall & Rossman, 1999). During the interpretation of the results it cannot be denied that there is the potential for observer bias, although the authors attempted to mitigate this by following a structured analytical process. Specifically, the authors relied on coding the transcripts using the NVivo software which makes it possible to view the individual steps undertaken in the analysis. Using such a software improves the replicability of the analysis, provided that the same data material is available.

Generalisability of the research design concerns itself with the extent to which findings are equally applicable to other environments, for instance other organizations (Saunders et al., 2009). In general, in qualitative research there is always an issue around generalisability due to the small and unrepresentative number of interview partners (Saunders et al., 2009). However, the authors actively took various decisions to increase the generalisability of the results. First, the authors chose representatives from companies that cover a number of industries (see Table 1). Second, interviews with consultancies were conducted in order to generate expert insights that could be used to validate the insights gained from the interviews with the product companies. Third, this study attempts to shed light on the generalisability of the results by analysing contingency factors that have an influence on the main relationship under observation. Thereby this study recognises the potential generalisability issues associated with the research design and has attempted to mitigate them where possible.

The *validity* of the research design describes whether the findings of the undertaken research are representing the research phenomenon in an accurate way (Collis & Hussey, 2003). In the context of this research, maturation is seen as a potential threat to the validity of the findings as it refers to the happening of events that might lead to a change in the researched phenomenon (Robson, 2002). As the field of BD is closely related to the developments in technology, ground-breaking changes in technology could lead to a change in the studied phenomenon. However, this is something that is in the nature of the study field and cannot be influenced by the authors.

3 Literature Review

In the following sections, the relevant literature underpinning the theoretical building blocks of the research questions is presented. Central concepts within these thematic areas are defined and related to each other. First, the literature in the area of NPD will be examined and different indicators to measure NPD performance discussed. Second, the field of BD, representing the overarching theoretical field in which BDA is placed, will be presented and its characteristics explained. Third, the literature on BDA is outlined and different types of BDA presented. Fourth, the emerging research field on the intersection of BDA and NPD is presented and central studies in this field introduced. Finally, the contingency factors which are hypothesized to influence the relationship between BDA and NPD will be motivated.

3.1 New Product Development

In order to fully grapple with the research question, the concept of NPD will be defined and placed within the wider innovation literature from which it obtains its theoretical foundation. The NPD process will then be explained in greater detail to highlight specific characteristics of each stage. Finally, the performance criteria typically used within NPD will be evaluated and discussed.

3.1.1 Definition of Innovation

In its simplest form innovation is *"the successful exploitation of new ideas"* (DTI, 2003, p. 8). Notably, the inclusion of the word 'exploitation' points to the fact that innovation is more than simply the generation of creative ideas, it is about the implementation of those ideas into a useful form to create economic value (Garcia & Calantone, 2002; Schilling, 2013). The OECD defines innovation as *"an iterative process initiated by the perception of a new market and/or new service opportunity for a technology-based invention which leads to development, production and marketing tasks striving for the commercial success of the invention."* (as cited in Garcia & Calantone, 2002, p. 112). Given that innovation is researched by a variety of scholastic communities, the associated literature can be categorised in a number of ways (R. Adams et al., 2006; Baregheh et al., 2009). One such division is according to the outcome of the innovation process, namely whether the innovation is related to either a product, service, process or business model (Baregheh et al., 2009; Foss & Saebi, 2017; Massa

& Tucci, 2014). As set out in the research question, this study will focus solely on product innovations, and the following subsection will define and explain the specific characteristics of this sub-field.

3.1.2 Product Innovation and the New Product Development Process

Although product innovation is an output, it also requires a process by which the innovation is conceptualised, developed and launched into the market, which is otherwise known as the NPD process. NPD involves *"converting an abstract idea into a tangible product, delivering it to potential customers when and where they want it, providing it at a price they are willing to pay, and earning at least a reasonable profit"* (Olson et al., 1995, p. 53). Somewhat surprisingly, there is little consensus in the academic community around a particular framework or theory for NPD (R. Adams et al., 2006; Veryzer Jr., 1998). Having said that, one of the more ubiquitous models for the NPD process is the Stage-Gate system, coined by Robert Cooper in the late 1980's (Cooper, 1990). The Stage-Gate system sets out *"a conceptual and operational map for moving new product projects from idea to launch and beyond - a blueprint for managing the new product development process"* (Cooper, 2008, p. 214). Cooper's (1993) initial Stage-Gate model encompassed the following stages: Idea discovery, Scoping, Business Case, Development, Testing and Launch. Although the Stage-Gate model is one of the more famous NPD models[1], what they all tend to have in common is the separation of the NPD process into discrete stages that are structured with quality controls and stop/go decisions (Jenkins et al., 2006).

Over time this lengthy and rigid system was adapted, both by Cooper and other academics within the field, to include fewer stages and allow for more iteration between stages (N. Bharadwaj, 2018; Cooper, 2014). A review of the more recent NPD literature reveals a trend towards an NPD process containing three overarching stages: Ideation, Product Development and Product Launch (Chang & Taylor, 2016; Ernst et al., 2010; Troy et al., 2008; Zhan et al., 2018).

The *Ideation* stage (also known as idea generation, concept development or fuzzy front-end phase) typically involves the *"generation and evaluation of new product ideas and further refinement of the most promising ideas into new product concepts"* which subsequently move into the development stage (Chang & Taylor, 2016; Ernst et al., 2010, p. 82; Zhan et al., 2018). The generation of ideas requires companies to essentially *"harness the voice of the customer"* so that they can obtain deep

---

[1] Others include; Phased Development, Product and Cycle-time Excellence and Total Design among others (Jenkins et al., 2006)

insights into user problems and generate ideas which can serve as the basis for potential product concepts (N. Bharadwaj et al., 2012; Griffin & Hauser, 1993, p. 1). To arrive at a final product concept the generated ideas are turned into visualisations which clarify the initial idea and allow for further refinement internally and with customers (N. Bharadwaj, 2018).

The *Product Development* stage is centred around taking the refined product concepts from the Ideation stage and developing them into tangible products (Ernst et al., 2010). More specifically, the Product Development stage consists of first developing the concept into a physical prototype, before testing it and utilizing feedback to arrive at the physical product (N. Bharadwaj, 2018).

The *Product Launch* stage (also known as the commercialisation stage) encompasses launching the product in the market and ensuring that it is a commercial success (R. Adams et al., 2006). Consequently, all tasks are directed either to inform customers about the new product or to ensure there is internal alignment about the new product such as through product training and sales support (N. Bharadwaj, 2018; Ernst et al., 2010). Given the nature of the tasks undertaken during this stage of the NPD process, marketing capabilities, as opposed to more technical capabilities, come to prominence (R. Adams et al., 2006)

3.1.3 NPD Performance

In order to answer the research questions, it is important to establish dependent variables, or metrics, by which it is possible to assess the influence of incorporating BDA on the NPD process. Therefore, the following subsection will outline, based on relevant literature, the different criteria used to evaluate an NPD process. In the conceptual framework, which follows the literature review, the key performance criteria applicable to this study will be motivated.

Unsurprisingly, given its importance to firm performance, there has been tremendous interest in the academic community in identifying the key factors which impact NPD performance, which, in the interest of clarity, refers to *"the success of new product development efforts"* (Troy et al., 2008, p. 136). However, the literature has failed to establish a definitive set of criteria upon which NPD performance should be judged (R. Adams et al., 2006). Leading academics within the field of NPD have proposed their own best practice NPD process (e.g. Cooper & Kleinschmidt, 1987) and then subsequently measured this process against selected NPD performance criteria to validate their theory. The issue arises due to the diversity of academic backgrounds that take an interest in NPD and innovation-related themes, resulting in equally diverse approaches to the measurement of NPD performance (R. Adams et al., 2006). A review of the meta-analyses which evaluate NPD success factors confirmed this diaspora of measures used as dependent variables with little coherent reasoning for using one over another (Ernst, 2002; Evanschitzky et al., 2012; Montoya-Weiss & Calantone, 1994).

Cordero (1990) provided an early summation of the different criteria to measure NPD performance that were being used in empirical studies and devised a classification for them that fundamentally distinguished between *"outputs and resources"* (p. 185). Outputs of the NPD process are the product(s) which are launched into the market and frequently used criteria are; the number of new products launched, the market share gained from new products or the payback period for a new product (Cordero, 1990). Resources represent the assets which are used in the NPD process (i.e. the inputs) and typical criteria identified include: (1) the financial spend incurred, (2) the number and quality of employees involved or (3) the time taken between the conception of an innovation and its launch into the marketplace (i.e. the duration of the NPD process) among others (Cordero, 1990).

Whilst Cordero (1990) identified that there needed to be measures for the inputs and the outputs, Adams et al. (2006) built on this work by identifying that there was an excessive focus on outcome-related criteria to judge NPD performance without considering criteria for the process itself. Consequently, Adams et al. (2006) consolidated the various measures used and created a comprehensive model which covers the traditional criteria, as presented by Cordero (1990), whilst also including additional criteria to assess the quality of the NPD process itself. In the following, the categories will be explained in detail.

Inputs Management reflects Cordero's (1990) concept of resources and refers to the management of the inputs to the NPD process. These inputs are defined as *"the raw materials or stimuli a system receives and processes, including people, equipment, facilities and funds"* (R. Adams et al., 2006, p. 27). In larger firms R&D intensity is typically used as a key metric highlighting the quantity of inputs provided into the NPD process. However, this is not a comprehensive measure as R&D is not the only input into the NPD process and it is of limited use for SMEs where formal R&D activities are lacking or not labelled as such (R. Adams et al., 2006). Authors have proposed using the level of funding as a proxy for inputs into the NPD process and/or the number of employees involved with the process.

However, both measures have been criticised as they only show the quantity of the input and not the quality, for example in terms of experience of the employees.

Commercialisation is similar to Cordero's (1990) concept of output and is concerned with making an innovative product a commercial success, which in turn, is essential for the survival and growth of organisations (R. Adams et al., 2006). Criteria in this category are frequently confined to quantitative measures such as the number of products launched in a given period, rate of adoption of a launched product or its financial performance.

In addition to these more traditional performance criteria, Adams et al. (2006) add extra categories which cover various aspects of the innovation process itself. *Knowledge Management* covers the management of explicit and implicit knowledge held by the organization (Nonaka, 1991). With regards to innovation there are three areas of importance: idea generation, knowledge storage (both implicit and explicit) and information flows (R. Adams et al., 2006). Attempts have been made to create quantifiable measures such as measuring the numbers of patents a company has or the number of linkages to external organisations that could provide knowledge.

Innovation strategy is generally understood to represent the articulation of a company's commitment to the development of products that are either new to the firm or the industry (R. Adams et al., 2006). As a measure of an NPD process, innovation strategy has two elements, firstly the strategic orientation, which is the extent to which the innovation strategy is embedded within the overall business strategy (Cooper, 1984). Secondly, is the strategic leadership which is important to provide a strong vision, long-term commitment, and a clear allocation of resources for innovation (Cooper, 1984).

Organisational culture and structure concern the way staff are grouped and the organisational culture within which they work (R. Adams et al., 2006). In line with the literature on organisational ambidexterity (e.g. March, 1991; Raisch & Birkinshaw, 2008) the organisational structure of a firm has an impact on its ability to innovate. A number of measures have been used in this category such as the degree of centralisation (i.e. concentration of decision-making authority) and formalisation (i.e. the extent to which rules and procedures are adhered to) of a firm (R. Adams et al., 2006). Within the organisational culture elements relating to motivation, propensity to take risks and shared vision of the company have been used as criteria to measure innovation performance.

Finally, Project Management is concerned with the processes that turn the inputs into marketable innovation outputs (R. Adams et al., 2006). A criteria frequently used for project management success is project duration or innovation speed, which has been positively correlated with product quality and fit with customer requirements (Hauser & Zettelmeyer, 1997).

3.2 Big Data

The research area of BD has received increasing interest from academia over the last few years (M. Chen et al., 2014). Scholars mentioned that BD might transform management theory as well as practice (George et al., 2014) and that it has the potential to radically improve firm performance (McAfee & Brynjolfsson, 2012). Agarwal & Dhar (2014) describe BD as the most important tech disruption in the field of business and academia since the emergence of the internet and the digital economy.

3.2.1 Five Vs of Big Data

Even though the importance of BD is widely recognized, a consistent definition of this abstract concept has not yet been established (M. Chen et al., 2014). Most of the diverse definitions of BD in literature and practice emphasize the characteristics of volume, velocity and variety (Mikalef et al., 2017). This 3V model was originally defined in a research report by Doug Laney (2001) and has been in widespread use ever since. Gandomi and Haider (2015) stress that there are no universal benchmarks for Big Data in terms of volume, velocity and variety, and that such defining thresholds depend on the sector, size and location of the company, and will evolve over time. Over the years, other researchers have added two further dimensions to the concept of BD, namely veracity and value. These five characteristics will be presented below:

Volume refers to the magnitude of data (Gandomi & Haider, 2015) and is described as being the primary attribute of BD (Russom, 2011). The volume of a data set influences in particular the degree to which the data can provide a comprehensive and granular view of the objects described in the data (Barton & Court, 2012). To describe the continuously increasing volume of data, prefixes such as peta- (1015), exa- (1018), zetta- (1021) and yottabytes (1024) were introduced (Bharadwaj, 2018). The World Economic Forum expects that by 2025 463 exabytes of data will be generated every day (Desjardins, 2019). A large share of this data will be created by social media. Already today roughly

800 million users on YouTube create an equivalent of 500 years of video per day and on Twitter 140 million active users post more than 340 million tweets (Fan & Gordon, 2014).The rapid growth in data volume will be further fostered by advances in information technology and IoT technology. IoT data, in particular, will account for the largest share of BD from 2030 onwards, as by then the number of sensors is expected to reach one trillion (M. Chen et al., 2014).

Variety describes the heterogeneity of the data in terms of data sources, formats and structure (Fosso Wamba et al., 2015; Russom, 2011). By increasing variety, data can give you a more diverse picture (Gupta & George, 2016) and companies that rely on data from a higher variety of sources will more likely stand out from competition (Zhao et al., 2014). In general, data can be classified as either internal or external data. Internal data represents all company-specific data that is generated by internal procedures and processes, whereas data collected from sources outside the company such as the internet, sensors or mobile phones are described as external data (Zhao et al., 2014). In terms of structural heterogeneity, data can be differentiated between structured data and unstructured data. Structured data comprise all tabular data, which can be found in spreadsheets or in relational databases. Unstructured data refers to text, images, audio and video, which sometimes are not structured enough to allow for data analysis. According to George et al. (2014) BD can also be categorized according to the respective source of data, as displayed in table 2.

Table 2: Classification of data according to data source (George et al., 2014; Gupta & George, 2016)

Public data	Public data is data owned by governments or local communities which can be used under certain restrictions and which is normally free of charge and usually includes topics like transportation or healthcare.
Private data	Private data is data held by private firms or individuals and thereby represents private information. Examples comprise data on customer transactions or data from mobile phone usage.
Data exhaust	Data exhaust can be described as ambient data. It is passively collected and has no direct value to the original data collector, but in combination with other data sources it might provide new insights. Examples are random internet searches or location data generated from mobile phones.
Community data	Community data is data that is generated by users in online social communities, such as Facebook and Twitter, or consumer reviews on products and can be described as unstructured.
Self-quantification data	Self-quantification data is data that is created by individuals through the quantification of personal behaviour and actions and can reveal insights about the psychology of those individuals. Examples are data generated from wearable technologies such as fitness bands and intelligent watches.

Velocity relates to the rate at which data is being generated and the speed at which it should be analysed and processed. High velocity of data enables companies to derive knowledge in a timely manner, which can be especially important in fast-moving environments (Johnson et al., 2017). The emergence and spread of digital devices such as smartphones and sensors has resulted in unprecedented data generation and a growing demand for real-time analysis and evidence-based planning (Gandomi & Haider, 2015).

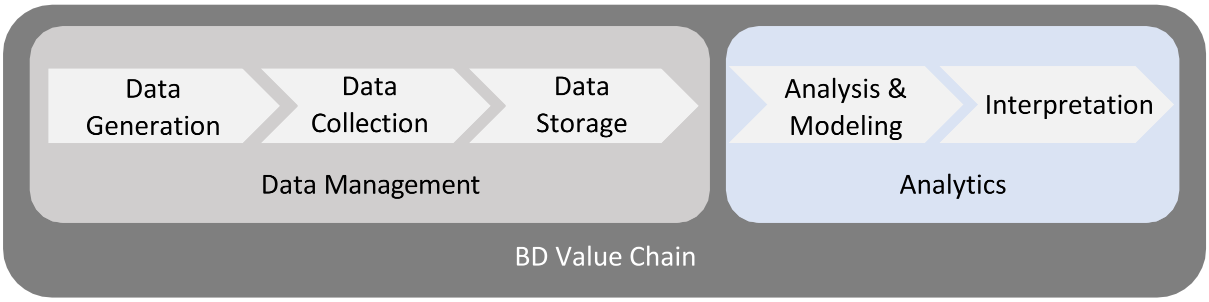
Veracity refers to the quality of the data in terms of completeness and reliability (N. Bharadwaj, 2018; Biehn, 2013). In particular, large data sets are quite often incomplete and do not contain data in all fields (Wedel & Kannan, 2016). Further, in other cases the gathered data is not fully reliable, which makes it necessary to assess the accuracy of the data (N. Bharadwaj, 2018). This dimension gained importance over time as one out of three managers do not trust the information used as the basis for decision making (LaValle et al., 2011).

Value describes the extent to which economically valuable insights and benefits can be derived from certain data (Fosso Wamba et al., 2015). In that context, Biehn (2013) highlights that having appropriate data will allow for more sophisticated queries and to analyse unique combinations within the data.

3.2.2 Value Chain of Big Data

The value chain of BD, which is depicted in Figure 4, can be generally divided into two phases, data management and analytics (Labrinidis & Jagadish, 2012). Within data management, it can be further differentiated between data generation, data acquisition and data storage (M. Chen et al., 2014). *Data generation* represents the first step and encompasses all generation of data, for example within companies or through IoT sensors. The second step is *data acquisition* and consists of data collection, data transmission and pre-processing. Within this phase data is collected and then transmitted to a certain data storage management system. The third step is *data storage*, which describes the storage and management of large-scale dataset. The second phase, analytics, can be distinguished between analysis & modelling and interpretation. *Analysis & modelling* comprises the methods and techniques to analyse BD and to obtain intelligence from it. Subsequently, in a second step called *interpretation*, the derived insights are verified and interpreted by a decision-maker (Labrinidis & Jagadish, 2012). In the following, this second phase of the BD value chain, which is also known as BDA, will be explained and the relevant literature outlined.

Figure 4: BD value chain (own illustration based on Chen et al., 2014 and Labrinidis & Jagadish, 2012)



3.3 Big Data Analytics

The value of BD unfolds only when the data is used for decision making. To make informed decision-making possible, companies need efficient processes to transform large amounts of fast-moving and diverse data into meaningful insights (Gandomi & Haider, 2015). In this context, BDA was described as a game changer allowing for better business efficiency and effectiveness in operations and strategy (Wamba et al., 2017) or as a crucial component of decision-making process in business (Hagel, 2015). BDA is defined as *"technologies (e.g. database and data mining tools) and techniques (e.g. analytical methods) that a company can employ to analyse large scale, complex data for various applications intended to augment firm performance in various dimensions"* (Kwon et al., 2014, p. 387). At the centre of BDA is the mining and extraction of patterns from enormous amounts of data for decisions, forecasts and other inferences (Najafabadi et al., 2015). The techniques that are used in this context include data mining, visualization, statistical analysis or design of experiments (in the form of A/B tests or multivariate testing) (Oussous et al., 2018). Moreover, new technologies and methods of analysis are applied to cope with the volume and speed of data and produce insights on an ongoing basis, for example using machine learning (Davenport, 2013).

3.3.1 Development of Big Data Analytics

The roots of general data analytics lie in the late 1960s, when practitioners started to analyse data with the use of computer systems to support decision making. Back then it was only performed in data-intensive business functions such as for analytical and repetitive tasks in areas like production planning, investment portfolio management, or transportation routing. In the 1970s statistical analysis performed via computer systems became more popular as computer applications were launched that made statistical analysis available (Davenport & Harris, 2007). According to Davenport & Harris (2007), since then the frontier of decisions for which data analytics can be valuable is constantly expanding. Business areas in which decisions were previously made based on intuition, accumulate more data and analytical rigor, and thereby can rely on data analytics instead of intuition.

Data Analytics has evolved drastically over time and, according to Davenport (2013), we have now entered the era of Data Analytics 3.0. In Data Analytics 1.0, coined *"the era of business intelligence"*, analysing data was consuming a lot of time and most of the analysis focused on reporting. Data analytics was focused on the past and contained no predictions or explanations (Davenport, 2013). By drastically increasing amounts of data and entering the era of Data Analytics 2.0 *("the era of Big Data"*) new powerful tools were needed. In this era, the data was no longer processed by a single server but shifted to more capable technologies that were able to cope with the increasingly unstructured nature of the data. Moreover, advances in machine-learning enabled handling of fast-moving streams of data. However, data analytics, whilst becoming much more powerful, were still mainly focused on back-office processes and tasks (Davenport, 2013). In contrast, Data Analytics 3.0 *("the era of data-enriched offerings"*) is characterized by the opportunity for every company, not just information firms and online companies, to also create products and services from their analysis of data. Data is now also analysed for the benefit of the customer and markets and so this new era has seen data analytics shift from the back office to the front-line of operations (Davenport, 2013). This is also highlighted by Mason (2018), who states that an excellent data strategy will not only comprise efficiency and performance improvements and automation, but also it will foster ideas and projects that generate new revenues sources and entire new businesses all derived from the company's data assets.

3.3.2 Types of Big Data Analytics

As outlined above, over the last decades both new types of data, such as self-quantification data, as well as new types of analysis have emerged. Hence, the status quo of BDA is characterized by a multitude of different types of data analytics, which can be classified broadly based on the time of execution, the type of data used for the analysis and the type of insights gained. The different types of BDA are presented below.

A first differentiation can be made between *real-time* or *offline* analysis. Whether the data is analysed in real time or offline, depends on the timeliness of the analysis. If the data is constantly changing, as for example in e-commerce, a fast data analysis makes sense to have the results available with as little delay as possible. If the reaction time is not that essential, then the data can be retrieved and analysed centrally in the company at a later point in time (M. Chen et al., 2014).

Another differentiation of data analytics types can be made based on the type of data used in the analysis. Besides the more traditional analysis of structured data, new types of analysis have emerged to also handle unstructured data (Gandomi & Haider, 2015). *Text analytics* can be used to derive information from textual data, which includes social network feeds, blogs, emails, online forums or survey responses. By conducting text analytics, large quantities of text can be transformed into insightful summaries, which can be used to support decision making. Within text analytics it can be differentiated between different methods. First, information extraction derive structured data from unstructured text (Gandomi & Haider, 2015). Second, text summarization can be used to derive the quintessence from one or several text documents. Third, question answering generates answers to questions formulated in natural language such as with Apple's Siri. Fourth, people's opinions towards products, certain events, individuals or organizations expressed in opinionated text can be derived through sentiment analysis (Gandomi & Haider, 2015).

Audio analytics refer to all types of analytics that derive information from unstructured audio data (Gandomi & Haider, 2015). Through audio summarization for instance, the most central words or also phrases from metadata can be derived, or a new representation of the audio content can be synthesized (M. Chen et al., 2014).

Video analytics is dedicated to extract insightful information from video streams (Gandomi & Haider, 2015). Thereby the most important or representative parts of video data can be interpreted (M. Chen et al., 2014). In retail for instance, video analytics can be used to analyse the shopping behaviour of different customer groups (Gandomi & Haider, 2015).

Social media analytics aims at extracting insights from social media data (Gandomi & Haider, 2015), which can be mainly classified as unstructured (Chan et al., 2016). Social media analytics can be split up into content-based analytics and structure-based analytics. The first one focuses on the content posted by users. For this purpose, text, audio and video analytics are applied depending on the type of post. The latter focuses on the structure of the network and looks at the relationships between the different users (Gandomi & Haider, 2015). For instance, with structure-based analytics, social graphs or activity graphs can be generated indicating either the relationships (e.g. friends on Facebook) or actual interactions between the users respectively (Heidemann et al., 2012).

Another differentiation can be made between descriptive, diagnostic, predictive and prescriptive analysis. First, *descriptive analytics* is used to describe a phenomenon by means of various measures that are able to cover its relevant aspects. It aims to disentangle what has happened (Banerjee et al., 2013). A common outcome of descriptive analytics is the identification of business opportunities and/or problems (Delen & Demirkan, 2013a). Second, *diagnostic analytics* can assess why something happened. This type of analytics involves exploratory data analysis of existing data or requires additional data that may need to be collected to determine the causes of a problem. Thereby it oftentimes relies on visualization techniques to accentuate the root cause(s) of a phenomenon (Banerjee et al., 2013). Third, *predictive analytics* seeks to answer what will happen and why it will happen (Delen & Demirkan, 2013b), thereby predicting potential future outcomes such as for instance future behaviour of a target customer segment (Banerjee et al., 2013). Therefore, either regression techniques or machine learning techniques (e.g. neural networks) are used (Gandomi & Haider, 2015). Finally, *prescriptive analytics* can be used to go beyond describing, diagnosing and predicting and answers what should be done in a particular situation to best achieve certain business objectives. This is done by linking different decision alternatives with the prediction of results that would occur in the case of these different alternatives (Banerjee et al., 2013).

3.4 Big Data Analytics for New Product Development

In the following section the literature on the intersection of BD or BDA and NPD is presented. In the interest of clarity, as BDA represents the second phase of the BD value chain, the potential of BDA is dependent on the data that is analysed. Therefore, literature looking at the relationship between BD and NPD is also included in this section, despite not being the construct under investigation. This is especially sensible, as the intersection between BD or BDA and NPD can still be regarded as being understudied (Zhan et al., 2017) and hence literature is rather scarce.

As mentioned by Davenport (2013), the importance and use of BD and the analysis of such large amounts of data for innovation has only increased dramatically in the last few years. Through the increasing digitalisation, companies are now operating in a data-rich environment and the analysis of BD can improve companies' innovation activities (Rindfleisch et al., 2017). This also applies to product innovations and scholars have highlighted that the increasing volume, velocity and variety of data is transforming the NPD process (Zhan et al., 2018). In line with this, Johnson et al. (2017) highlights that the assimilation between product innovation and BD can be found in many industries and will lead to major changes in the NPD process.

The integration of BD into the NPD process can in turn provide various advantages. Through the analysis of BD, companies can derive information about their customers and can understand their preferences better (Xu et al., 2016), which enables companies to develop more customer-centred products (Zhan et al., 2018). Further, through the use of BD companies can leverage market opportunities before they become obvious (Johnson et al., 2017). Moreover, the analysis of BD will enable companies to improve the testing of prototypes and to acquire feedback (Xu et al., 2016). Manyika et al. (2011) exemplified the ability of BDA to improve the development process of products, stating that pharmaceutical companies can reduce the average time needed to develop a new drug by three to five years through the use of BDA. Another concrete example of the value of BD in the NPD process was given by Xu et al. (2016), who outlined that the American company Netflix analyses millions of real-time data points generated by its viewers to ascertain whether a pilot will succeed as a new show. Moreover, the company conducts large scale experiments including test and control groups to better understand the perceptions of consumers with regards to new products. Hence, BDA is leveraged to create original video content and to make multi-million-dollar decisions about new products.

Despite these statements affirming the potential of BD or BDA for the NPD process, only a few scholars have conducted studies which specifically research the use of BD or BDA for NPD or related activities and topics. These studies are outlined in table 3 and are classified according to their focus on BD or BDA.

Table 3: Key studies in the intersection of BD/BDA and NPD

Author (Year)	**Research focus**
Studies looking at BD and NPD (or related themes)	
Johnson et al. (2017)	This study empirically examined what effect a company's exploitation or exploration orientation has on the three characteristics (volume, velocity and variety) of BD usage. Only an exploration orientation was found to have an effect on all three. When customer demands are very volatile the importance of data velocity for the success of new products is higher, while the importance of volume is lower.
Zhan et al. (2017)	This study identified key success factors in a BD environment, that can foster better planning and organisation of the innovation process, through which the speed and cost of the NPD process could be improved.
Zhan et al. (2018)	This study investigated how through using BD companies can identify unrecognized needs. The authors developed a customer involvement approach for NPD to utilise customer information.
Studies looking at BDA and NPD	
Chan et al. (2016)	In this study the authors developed a mixed method BDA approach to leverage social media data for decision making in NPD. The study highlighted the economic benefits of the approach, as social media data is freely available online and thereby much cheaper than collecting such information by other means.
Xu et al. (2016)	This study researched the relationships between BDA, traditional marketing analytics and new product success. The study emphasized that more traditional activities, such as consumer surveys, can still be used in the NPD process in some situations, but in volatile and fast-moving industries and markets NPD is much more complex and hence large amounts of fast-moving data from a variety of sources are needed to understand both customers and markets properly.

In addition to the studies outlined above, limited research can also be found on the use of BDA in all three phases of the NPD process. The respective literature is presented in the order of the NPD process, starting with the Ideation stage.

With respect to the *Ideation* stage, different scholars have researched certain potential use cases for leveraging BD for the generation and evaluation of new ideas and the development of first product concepts. The majority of these studies have focussed on the ability to generate ideas from Social Media although each one has taken a particular angle. Lee and Bradlow (2011) found out, that product attributes and their relative importance for customers can be derived through social media analytics, for instance from ratings and review data. Through social media analytics the potential openings in the market can be identified by analysing the competitive landscape (Netzer et al., 2012). Liu and Kop (2016) found that the incorporation of social media data will allow firms to minimise uncertainties as they will better understand customers' currently unmet needs and their expectations towards new products. In particular, their study outlines that social media data can improve the quality and variety of ideas for new products. However, their study also emphasizes the challenges to make sense of these enormous amounts of data, of which a lot will be rather irrelevant information. In a study conducted by Tuarob and Tucker (2015) the authors developed a data mining driven methodology to identify lead users on social media platforms. Through a case study the applicability of the approach was subsequently proven. The approach can both automatically identify lead users as well as potential latent product features proposed by those lead users, which can be used for the development of new products or product refinement (Tuarob & Tucker, 2015). Moreover, Christensen et al. (2017) outline that, through machine learning and text mining, ideas can be automatically identified in online communities, which is especially valuable as such communities often contain vast amounts of information, much of which is not valuable and does not contain concrete ideas. Similarly, Hoornaert et al. (2017) highlight that through data analytics companies can automatically detect the best ideas from crowdsourcing communities based on algorithms analysing the content, the contributor and the crowd's feedback.

With respect to the *Product Development* stage, the literature can be described as being less exhaustive. Scholars have mentioned the potential of the use of BD and BDA for the development of a first prototype, subsequently testing and refinement of the prototype until arriving at a physical product. Jin et al. (2016) researched how BDA on consumer opinions can be used to inform market-driven product design. The identification of consumer requirements towards the new product and a data analysis of competitor products were used to inform the designer about the pros and cons of different product designs, thereby enabling better product design. Similarly, Li et al. (2015) highlighted that BD can be beneficial in the whole design process, by first deriving specific functions for the design, then defining solutions that match the design specifications and finally, in the decision-making around the final details of the product design. Moreover and with regards to testing, Fan and Gordon (2014) highlight that especially for the software industry, social media-based testing is used to release various versions and then gather feedback on those versions.

With respect to the *Product Launch* stage literature is barely existent. On the one hand, Moe and Schweidel (2017) outline that marketing insights can be derived through social media analytics, which can help to place the product effectively in the market. On the other hand, Baker et al. (2014) state that BD can be used to derive more suitable prices on much more granular product levels. The analysed companies were thereby able to increase their profit margin between three and eight percent.

In summary, it can be identified that more and more scholars are outlining the potential of using BD in the NPD process and that these authors are emphasizing that such an integration can provide diverse benefits for the respective companies. Studies published so far can be split up into two areas. On the one hand, a few studies have explored the relationship between BD/BDA and NPD related themes such as exploration and exploitation activities of firms. On the other hand, concrete examples on how to leverage BD or BDA respectively in the NPD process for specific activities are presented. However, these studies mainly focus on the Ideation stage. Notable, there is a lack of literature within the research area of BDA related to Product Launch. Given the already highlighted point that the Product Launch involves more marketing related capabilities, it might be the case that there is more extensive literature touching upon activities that are relevant to this stage in the general marketing research area.

3.5 Contingency Factors of BDA-NPD Performance Relationship

In addition to exploring the relationship between BDA and NPD performance it is also useful, both from an academic and practitioner perspective, to highlight those contingency factors that may influence the strength of this relationship. The range of potential contingency factors to investigate is daunting, due to the numerous papers which tackle this issue from an NPD perspective (e.g. Calantone et al., 2010; Ernst, 2002; Ernst et al., 2010; Troy et al., 2008 among others) and conversely due to a lack of literature in the intersection of BDA and NPD to narrow the scope of possible contingency factors (Mikalef, Pappas, et al., 2018). As previously outlined, this thesis will focus on the level of organisational agility and the degree of innovativeness of the product innovation. In the following subsection, both contingency factors will be described and the current linkages that exist in the literature will be explained.

3.5.1 Organisational Agility

In addition to the degree of product innovativeness another contingency factor appears important, namely, organisational agility. Organisational agility is derived from the broader word agility, which is defined as *"the ability to move about quickly and easily"* (Cambridge Dictionary, 2019). Within the field of management, both practice and academia, the idea of being able to move quickly and easily has gained importance as companies face increasingly volatile environments (Johnson et al., 2017; Pavlou & El Sawy, 2011; Sambamurthy et al., 2003). Such environmental changes led to the development of a new theoretical concept called *dynamic capabilities* which refers to *"the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments"* (Teece et al., 1997, p. 516). Dynamic capabilities encompass a variety of sub-capabilities particularly related to sensing, learning, integrating and coordinating, and organisational agility is cited as an example of a dynamic capability (Pavlou & El Sawy, 2011; Sambamurthy et al., 2003). However, the dynamic capabilities field has been criticised for a *"lack of precise definition, empirical grounding, and measurement"* (Pavlou & El Sawy, 2011, p. 240) and consequently organisational agility, being considered a type of dynamic capability, suffers from similarly loose definitions. Some authors have referred to this ability as a firm's 'responsiveness' (Mikalef et al., 2019), others call it a 'reconfiguration ability' (McCardle et al., 2018), whilst others define it in broader terms such as 'process-oriented dynamic capabilities' (Wamba et al., 2017). However, whilst the exact name given to the capability can vary the underlying notion is largely the same and in the information systems literature the term *organisational agility* dominates (Shuradze et al., 2018).

Sambamurthy et al. (2003, p. 245) define organisational agility as a company's ability *"to detect opportunities for innovation and seize those competitive market opportunities by assembling requisite assets, knowledge, and relationships with speed and surprise"*. Pavlou & El Sawy (2011, p. 260) define organisational agility as the *"ability to sense and respond"* and they place it within the broader dynamic capabilities concept by saying that it more closely resembles the *"respond component"* of dynamic capabilities. Sambamurthy et al. (2003, p. 238) concur that organisational agility is an example of a dynamic capability and further note that they do not impact firm performance directly, rather they achieve this by *"impacting the quality of competitive actions by firms"*. Common examples of such competitive actions are moving into a new market segment or launching a new product into the market (Sambamurthy et al., 2003). Put succinctly, organisational agility, by being an example of a dynamic capability, has the potential to influence the quality of competitive actions, such as launching a product. Therefore, the theoretical foundation of organisational agility within dynamic capabilities suggests that it may influence the relationship between BDA and NPD performance.

Davenport et al. (2012, p. 46), identify the need for organisational agility when they suggest that a *"key tenet of big data is that the world, and the data that describe it, are constantly changing, and* *organizations that can recognize the changes and react quickly and intelligently will have the upper hand"*. This need for organisational agility is not exclusively related to NPD and there is some literature covering organisational agility and BDA in various functional areas of the firm, notably supply chain management (e.g. Dubey et al., 2018; Giannakis & Louis, 2016; Nguyen et al., 2018). Wamba et al. (2017) seem to validate the assertion by Davenport et al. (2012) as they find that process-oriented dynamic capabilities, a term similar to organisational agility, play a mediating role between BDA and overall firm performance.

However, there is very limited literature which empirically examines the impact of organisational agility on NPD performance. A study by Shuradze et al. (2018) provides initial empirical evidence in support for organisational agility influencing innovation success. Their study finds that organisational agility acts as a partial mediating variable in the relationship between marketing-enabled data analytics capability (MDAC) and innovation success (Shuradze et al., 2018). Shuradze et al. (2018) include all forms of innovation in their scope and define innovation success according to two dimensions depending on whether the innovation is explorative or exploitative. The independent variable, MDAC, is rooted in the IT capabilities literature and as such has some similar elements to the BDA concept however they are far from being the same (A. S. Bharadwaj, 2000; Shuradze et al., 2018). A key difference is that MDAC is defined as a firm capability to deploy analytics and marketing resources in combination, and as such is more similar to the concept of BDAC described above (Gupta & George, 2016; Shuradze et al., 2018). In contrast, BDA has no specific functional focus but comprises the sum of technologies and techniques that can be used to analyse BD. While the Shuradze et al. (2018) study set out to empirically establish the link between organisational agility and innovation success there is little guidance as to how this relationship functions in reality. Therefore, from their study it is hard to ascertain why organisational agility is so important for innovation success and what specific elements of innovation success are affected.

In summary, the theoretical basis of organisational agility as a dynamic capability suggests that it could have an impact on a companies' NPD process. Literature linking BDA and organisational agility has mainly focussed on supply chain management and the limited papers focussing on NPD do not explore how organisational can influence this process nor what aspects of it are affected.

3.5.2 Product Innovativeness

As outlined, innovations can be categorised according to whether the outcome of the process is either a product, service, process or business model, but within those categories it can be further subdivided according to the innovativeness of the innovation (Garcia & Calantone, 2002). There has been little agreement with regards to how to define the innovativeness of a product (Henard & Szymanski, 2003; Henderson & Clark, 1990), although Garcia and Calantone (2002) have attempted, through a comprehensive literature review, to define this term. According to Garcia and Calantone (2002) product innovativeness is *"a measure of the potential discontinuity a product can generate in the marketing and/or technological process"* (p. 113). Furthermore, this discontinuity can either be on a macro or micro level; macro creating a paradigm shift in science or new market structure in an industry, whilst micro is a discontinuity at the level of the firm (Garcia & Calantone, 2002). According to this distinction between first-order (macro/micro) and second-order (market/technology) levels, it is possible to categorise and assign a typology to the level of innovativeness of a product.

At one end of the innovativeness spectrum are radical innovations, which are defined as *"innovations that embody a new technology that results in a new market infrastructure"* (Garcia & Calantone, 2002, p. 120). In other words, radical innovations introduce technological and market discontinuities on a macro level, creating new demand previously unrecognised by the consumer and subsequently new industries, competitors, distribution channels and marketing activities (Song & Montoya-Weiss, 1998). In contrast, at the other end of the scale incremental innovations merely *"provide new features, benefits or improvements to the existing technology in the existing market."* (Garcia & Calantone, 2002). Therefore, incremental innovations do not involve any macro discontinuities and just create a discontinuity at the level of the firm. Between the two ends of the spectrum, Garcia & Calantone (2002) identify a middle category they term as 'really new' innovations, where there is <u>*either*</u> a new technology <u>*or*</u> a new market created by an innovation at the macro level, but not both, as in the case of radical innovations.

Studies in the NPD literature suggest that the NPD process differs according to the level of innovativeness of a product (Garcia & Calantone, 2002; Reid & De Brentani, 2004; Song & Montoya-Weiss, 1998; Veryzer Jr., 1998). The NPD process is affected by the level of technological uncertainty in radical innovations as it can distort the progression through the NPD process (Chesbrough et al., 2006). Innovations containing new technology can require longer periods spent in the Product Development phase as the technology develops, and more iteration between the Ideation and

Product Development stages as technological changes are incorporated (Veryzer Jr., 1998). Further, new technology can negatively impact the diversity of innovation teams, due to the more specific technological knowledge domain required (Veryzer Jr., 1998).

Moreover, the more exploratory nature of radical innovations, both from a technology and market perspective, creates a NPD process that is less customer driven than for more incremental product innovations (Garcia & Calantone, 2002). The rationale for this is twofold. First, it can be difficult to identify who the radical innovation is for as it involves new technology targeted at a new market (Veryzer Jr., 1998). Secondly, radical innovations are also more distant from the current market and so customers are inherently less familiar with the innovations, limiting their ability to provide needs-related input or to test product ideas (Liu & Kop, 2016). In addition, the secrecy involved with the development of new technologies also impacts the willingness to integrate customers, due to the threat of knowledge leakage (Veryzer Jr., 1998).

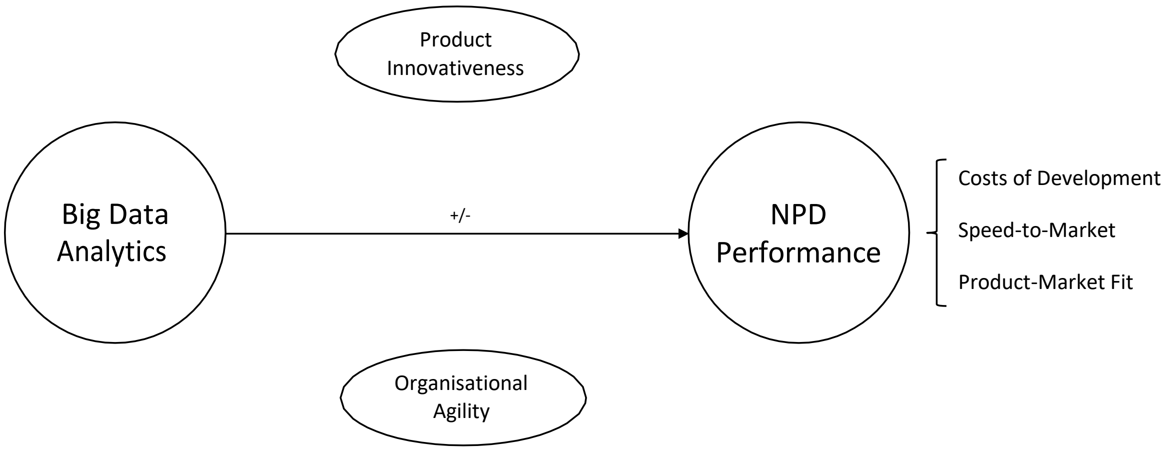
Finally, the result of increased uncertainty associated with radical innovations is that the decision-making within the NPD process differs from incremental innovations (N. Bharadwaj, 2018). Whereas *"decisions and activities tend to be more explicit and structured"* for incremental innovations (Reid & De Brentani, 2004, p. 172), the high-degree of uncertainty and iteration involved in radical innovations creates an inherently chaotic process (Veryzer Jr., 1998). A potential consequence of this less structured process is the finding that managers suffer from decision-making biases, for example they are more likely to persevere with a NPD project the more innovative it is considered (Garcia & Calantone, 2002).

Section 3.4 highlighted that the literature which covers the intersection between BDA and NPD is sparse given the emerging nature of the research area, and so those papers which specifically include product innovativeness as a variable are even more so. In the wider BD research field, some papers have indicated a potential relationship between BD and the degree of innovativeness. While for instance Zhan et al. (2018, p. 592) mentioned that BD plays an *"important role in enabling companies to come up with genuinely innovative new products"*, Hartmann et al. (2016, p. 1385) suggested it can be used for incremental improvements and optimisations. The only paper which takes a similar approach to this thesis is a study by Mikalef et al. (2019) which researched the extent to which a company's BDA capability enhances both incremental and radical innovation capabilities. The study found that BDA has a positive effect on both, incremental and radical innovation capabilities (Mikalef et al., 2019). However, while this paper empirically proves a link between BDA and product innovativeness there is still a lack of research and guidance as to how it affects the use and value of BDA for NPD. Bharadwaj (2018) notes that practitioners in NPD would benefit greatly from guidance on to how increase the success of product innovations with different levels of innovativeness and calls for further research to fill this void.

4 Conceptual Framework

The conceptual framework was developed based on both the initial data collection phase, which narrowed down the scope, and on the pertinent literature outlined above. The conceptual framework provides a clear and succinct understanding and, where appropriate, an operationalisation of the terms under investigation. As outlined by Saunders et al. (2009) developing such a conceptual framework, is crucial to guide the further data collection and analysis in studies with an exploratory purpose. Figure 5 below shows the conceptual framework.

Figure 5: Conceptual framework (own illustration)



4.1 Definition of Big Data Analytics

In line with the above stated definition of BDA by Kwon et al. (2014), the authors understand BDA as the technology and techniques that can be leveraged by companies to analyse BD in order to improve the performance of the firm. As this thesis aims to explore the potential of BDA for NPD broadly, all types of BDA, as presented in section 3.3.2, are included with no specific focus on certain technologies or techniques. Moreover, while the integration of BDA into business areas that are not related to NPD could in some way also affect NPD performance, this study deals solely with BDA insofar as it is integrated into the NPD process.

4.2 Definition of New Product Development Performance

As discussed above there is a diaspora of criteria used to evaluate NPD performance. However, this thesis will consider three overarching criteria which sequentially cover the entire NPD process. In line with many previous studies covering NPD, the overall costs of development will be the criteria to assess the inputs to the NPD process (R. Adams et al., 2006; Cordero, 1990). The criteria to evaluate the NPD process itself will be the speed-to-market of the product innovation. Finally, for the output of the NPD process the designated criteria is the product-market fit. Each of these three criteria will be further explained below as will the underlying rationale behind their selection.

With regards to the *Costs of Development* criteria, Adams et al. (2006) define the inputs to the NPD process as the raw materials a system receives and processes, which can include people, equipment, facilities and funds. Whilst this definition distinguishes between the different types of inputs to the NPD process, fundamentally all of the inputs are translated into monetary terms at some point, when either the equipment is purchased, or the people hired. Consequently, this thesis will use the general measure of costs associated with the NPD process as the input criteria.

Although Adams et al. (2006) propose a number of different criteria to evaluate the quality of the NPD process, such as knowledge management, it is beyond the scope of this study to be able to consider all of them due to the complexity and time needed to undertake such a challenge. As a result, it was necessary to select a single measure which would reflect, to some extent, the NPD process itself. New product speed-to-market depends on how quickly a firm can handle issues of coordination and control across each stage of the NPD process (Johnson et al., 2017). Therefore, if a firm is able to achieve a high speed-to-market it must have controlled, to some extent, the issues of coordination and control. Consequently, it is assumed that achieving a high speed-to-market is a partial reflection of the NPD process itself. Furthermore, whilst a faster speed-to-market has an impact on development costs by reducing the duration of the NPD process, it can also provide benefits which can aid the success rate of a product. Schilling (2013) notes that a faster speed to market can also lead to products being 'first-movers', which can, at least for a limited time, lead to technological leadership, the pre-emption of scarce assets by the first mover or the creation of buyer switching costs (Lieberman & Montgomery, 1988). Therefore, this thesis will use *Speed-to-Market* to evaluate the effectiveness of the NPD process itself.

Finally, focussing on *Product-Market Fit,* Adams et al. (2006) note that many studies have used outcome criteria to evaluate NPD performance, typically a financial measure such as return on NPD investment. However, as the costs of the NPD process have been accounted for separately, the criteria to evaluate the output should measure the success of the product which is launched onto the market. For a product to be successful on the market, it must be desired by a certain customer segment. To achieve this the product must satisfy that segment's given product requirements, whether these requirements are known or not by the customers. These requirements could be met by, for example, offering more compelling features, greater quality or more attractive pricing than competing products (Schilling, 2013). Ensuring that the product meets the requirements of the market is also known as creating a *product-market fit*, which has been shown to lead to new product financial success (i.e. new product revenue, sales or market share) (Johnson et al., 2017). Therefore, product-market fit will be used throughout the analysis as it is the criteria which best evaluates the output of the NPD process without including cost criteria.

4.3 Definition of Organisational Contingency Factors

4.3.1 Organisational Agility

Given that BDA is a term that is embedded within the information systems research field, this study will employ the use of the term organisational agility. In line with Sambamurthy et al. (2003, p. 245), this study defines organisational agility as a company's ability *"to detect opportunities for innovation and seize those competitive market opportunities by assembling requisite assets, knowledge, and relationships with speed and surprise"*. Furthermore, this thesis also recognises organisational agility as a dynamic capability, and as such it is one that is able to influence the quality of competitive actions, such as launching a new product into the market.

In order to effectively measure the influence of this concept on the relationship between BDA and NPD performance it must be operationalised. As stated in the introduction, this study aims to explore how the relationship between BDA and NPD performance is contingent upon the level of organisational agility. Therefore, as the emphasis is on *how* it is contingent, interviewees were asked open questions instead of using rigid guides, in order to allow the respondents to reflect on the nuances of how organisational agility influences the main relationship between BDA and NPD performance. In contrast to Shuradze et al.'s (2018) operationalisation of organisational agility, this study also favours the holistic use of the term organisational agility as this allows for unexpected answers from the respondents based on their own interpretation and experience from their respective companies. Responses from interviewees from product companies are assumed to reflect their respective companies whilst responses from consultants are assumed to be without company or industry bias.

4.3.2 Product Innovativeness

Garcia and Calantone (2002) provide a comprehensive review of the literature around product innovativeness and set out the three categories, or levels, of innovativeness a product can fit into. Whilst the middle category of 'really new' innovations is a valuable addition to the innovation literature it adds unnecessary complexity to this study, which is only interested in the relative difference in product innovativeness and how this could impact the quality and value of the insights generated from BDA which can influence NPD performance. For example, this study is only interested in whether product X is more radical than product Y, not which exact innovativeness category product X should fall into. Therefore, this study will classify innovations on a spectrum with radical at one end and incremental at the other, although the definitions of each of these terms will be in accordance to that set out by Garcia and Calantone (2002) in their study. Furthermore, to aid with the subsequent analysis it is assumed that there is a linear relationship between incremental and radical innovations and as such any relationship highlighted in the results will also be assumed to be linear.

5 Results & Analysis

The results of the analysis have been divided into four sections which follow the order of the research questions outlined in the introduction. Section 5.1 will present a typology for using BDA in the NPD process. Section 5.2 will outline the results which focus on how the incorporation of BDA influences NPD performance. The results regarding the two organisational contingency factors will be presented in section 5.3. Section 5.4 will conclude the results and analysis chapter by presenting the *ex-post* results.

5.1 Typology of Big Data Analytics for the New Product Development Process

In general, all 12 interview partners confirmed that there is great potential for companies to incorporate BDA into the NPD process. In fact, Shah (2019) even went so far as to say that, in his opinion, *"it's a very obvious answer […] I would really be surprised if somebody says no there is no link"*. Further, Ahlbrand was underlining the potential by saying that during the *"last 3 years it [BDA in NPD] has increased by 100% from year to year. In any case, this will continue for the next few years. I also don't think that this department is getting smaller, but that more and more people are realizing that this is important"*. However, eight respondents suggested that besides the generally positive relationship, there are also caveats to consider. Further, despite of the great potential of using BDA for NPD, the interview partners also acknowledged that the current use of BDA in companies is mainly focused on operational issues, namely operational efficiency or optimisation. Hede (2019) suggested that this is most likely due to the fact that creating investment cases for these types of projects *"is just super easy"* and therefore companies tend to focus on that area more than NPD. However, in contrast to his previous statement, Hede (2019) did allude that whilst the use of BDA in operations is easier the *"really big wins are going to be in new products and services"* which adds credence to the importance of the current study.

Table 4: Interview respondent analysis

Name	Role	Ideation	Product Development	Product Launch
Kiran Vas	Consultant	Positive effect and use case	Positive effect and use case	Positive effect
Adam Hede	Consultant	Positive effect and use case	Positive effect and use case	Positive effect and use case
Abayomi Baiyere	Academic	Positive effect and use case	Positive effect	Positive effect
Mathias Blom	Consultant	Positive effect and use case	Positive effect and use case	Positive effect and use case
Asbjørn Andersen	Consultant	Positive effect and use case	Positive effect	Positive effect
Rahul Shah	Consultant	Positive effect and use case	Positive effect and use case	No statement
Isa Anjos	Product Company	Positive effect and use case	Negative effect	Positive effect
Patrick Ahlbrand	Product Company	Positive effect and use case	Positive effect and use case	Positive effect and use case
Hannah Thomsen	Product Company	Positive effect and use case	Positive effect and use case	Positive effect and use case
Sebastian Barfort	Consultant	Positive effect and use case	Positive effect and use case	Positive effect and use case
Nicolas Antille	Product Company	Positive effect and use case	Positive effect and use case	No statement
Nils Dulfer	Consultant	Positive effect and use case	Positive effect and use case	Positive effect and use case

Positive effect and use case   Positive effect   Negative effect   No statement

The following section will present a typology for using BDA in the NPD process. An overview about this typology is presented on the next page (Figure 6). In the conducted interviews, 26 use cases were described by the respondents, which, after several rounds of analysis, could be grouped into 11 overarching themes. These themes represent the potential means by which BDA can be incorporated into Ideation, Product Development and Product Launch. Out of the 11 themes, four of them fall into the Ideation stage, namely *Trend Spotting*, *Customer Preference Identification*, *Input Scanning* and *Quality Improvement*. Further, four themes could be identified in the Product Development stage which are *Customer Preference Matching*, *Performance Optimization*, *New Product Exploration* and *New Product Verification*. Finally, the three themes in the Product Launch stage are *Promotion Tailoring*, *Price Determination* and *Launch Calibration*. In the following, the themes will be explained in the order in which they would occur in the NPD process. Within each theme the specific use cases provided by interviewees will be outlined to concretize the themes and show how they are applied in practice.

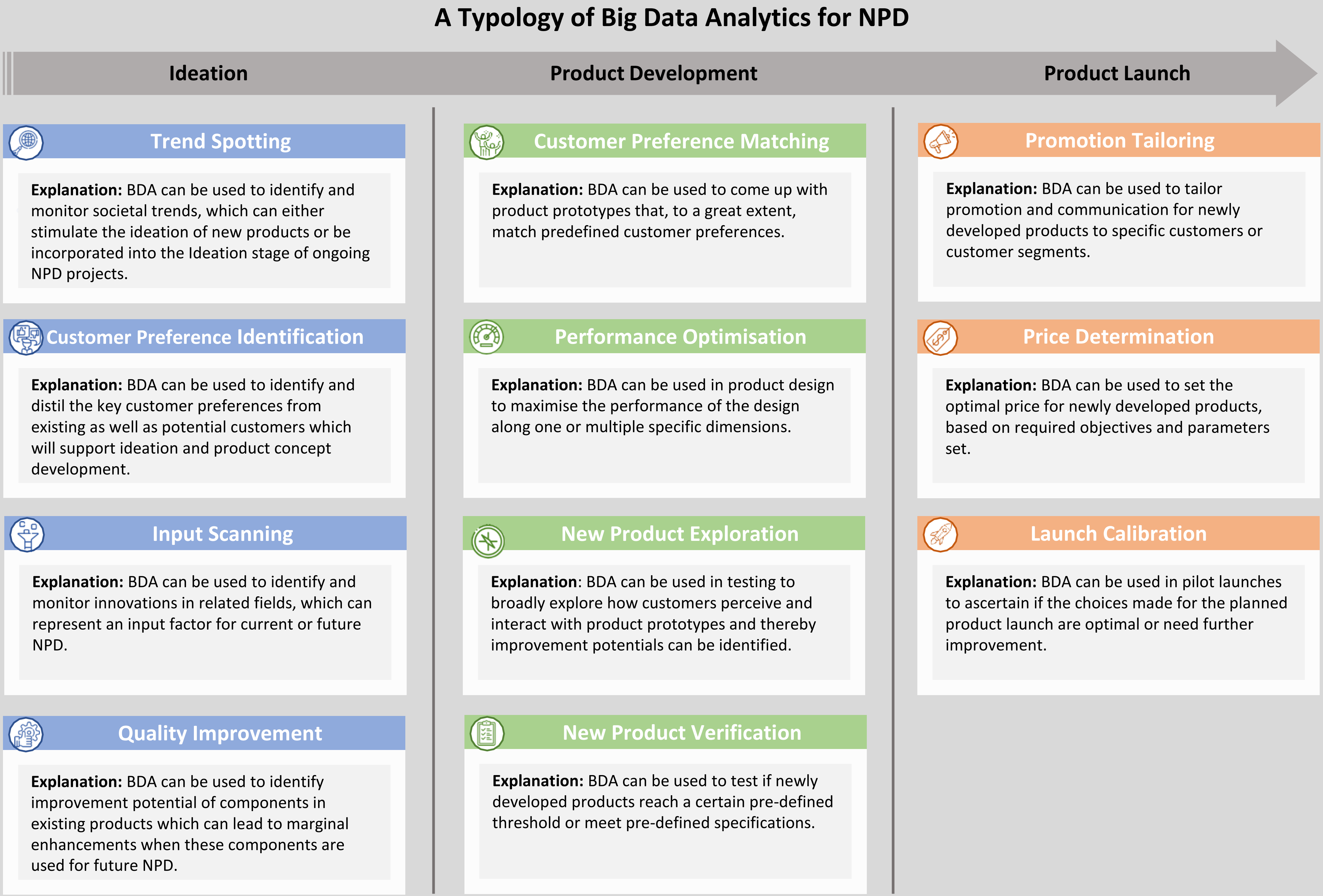


Figure 6: Overview of the typology of BDA for NPD (own illustration)

5.1.1 Big Data Analytics in the Ideation Stage

As outlined in the literature review, this thesis considers the Ideation stage to involve both the generation and the evaluation of ideas which are then refined into new product concepts ready to be moved into the development phase (Ernst et al., 2010). With this definition in mind, all 12 interview partners highlighted the positive potential of incorporating BDA for the Ideation stage and eight of them were also able to provide a specific use case where BDA is already being used in practice (Table 4).

As an example of the positive potential, Dülfer (2019) stated that *"at the moment the biggest chances [of using BDA] are probably with regard to the very early front end"*. The main reason for this is due to *"all the information that is available from customers"* which allows companies to *"identify early on preferences in order to start the Ideation process"* (Dülfer, 2019). Vas (2019) confirmed the importance of data in helping companies understand their customers' needs and problems, stating that the minute he receives the data about the customer he knows their problems. Thomsen (2019) extended this view further and suggests that the ability to be able to *"incorporate different sources of data"* is key in the Ideation phase as it allows a company to be *"less reliant on just our own history"* and therefore avoid repeating just the same product ideas over again which *"at one point becomes boring"*.

Despite generally being positive about the potential of incorporating BDA in the Ideation stage, Barfort (2019) expressed some concern saying that he was *"sceptical as to how data analytics is informative"* for coming up with initial ideas or insights. Baiyere (2019) provided a possible explanation for this limitation of BDA in the Ideation phase when he suggested that because *"data will always be based on history"* then it makes it difficult to *"imagine different scenarios"*. Consequently, Baiyere (2019) believes that this provides a delineation between what is and is not possible by using BDA in the Ideation phase. In contrast, Hede (2019) noted that *"we are taking a serious crack at creativity within machine learning"* due to the continuous advances in technology, but still acknowledges that this creativity is still based on *"taking averages of things"* and as such may be limited.

From the insights and perspectives gathered from the interviews it was possible to identify 4 themes that represent means by which BDA can be incorporated in the Ideation stage. The four themes identified are: *Trend Spotting, Customer Preference Identification, Input Scanning and Quality*

Improvement. The more exploratory nature of *Trend Spotting* and *Customer Preference Identification* means that they will be explained first, followed by *Input Scanning* and *Quality Improvement,* which are more closely related to the Product Development stage.

5.1.1.1 Trend Spotting

The *Trend Spotting* theme highlights how BDA can be used to identify and monitor societal trends. Four interviewees suggested that BDA could be used in this regard (Barfort, 2019; Blom, 2019; Dülfer, 2019; Thomsen, 2019). The mentioned data leveraged for this theme can be classified as external data, mainly representing community data from social media platforms, but also data exhaust such as data generated through web searches. This kind of data is freely available and can be analysed in real time. The type of analytics pursued in this theme is mainly social media analytics. Two specific use cases fall into this theme. The first use case shows how BDA can be incorporated to identify current or upcoming trends and the second use case demonstrates how people's sentiment can be gauged towards certain societal trends (Barfort, 2019; Blom, 2019; Dülfer, 2019; Thomsen, 2019).

The first use case was proposed by both Thomsen (2019) and Dülfer (2019) and highlights the potential of BDA to identify current or upcoming trends that are relevant for the company's customer segments. According to Dülfer (2019), more and more companies are interested in a so-called *"trend radars"* which, based on BDA, can identify relevant trends in society. While both interview partners highlighted the value of community data to identify what people are looking for, Thomsen (2019) also mentioned the value of web search data. These examples of external data are able to provide insights into trending topics that have historically not been explored by the company (Thomsen, 2019). Dülfer (2019) provided a specific explanation of this analysis and outlined that companies can run cluster analyses on social media data, especially Twitter. Essentially, companies can detect if people start tweeting something that is very unusual against a certain baseline. By defining certain trigger words, companies can identify when these trigger words are suddenly used in combination with a new word. Dülfer (2019) exemplifies that by stating that the trigger word *"delicious"* could be used to identify when customers increasingly express their liking for a new type of food. The analytics would then inform the company about these upcoming trends either in real-time or compiled in daily, weekly or monthly updates.

The benefits that such BDA has was highlighted by both interview partners through concrete examples. Thomsen (2019) mentioned that the increasing interest in climate related issues represented an example of such a trend that had not previously been considered, which could subsequently be incorporated into the Ideation of new products. Therefore, as Thomsen (2019) puts it *"just looking at our own history will not find it"*. Whilst this type of information could be obtained through other external sources, such as traditional media coverage, Thomsen (2019) argued that using BDA allows for a quantification and filtering of the most prevalent trends in society. *"Here it does matter if you know 5,000 people have searched for horses but 10,000 have searched for unicorns"* (Thomsen 2019). Fundamentally, it was stressed that in Pandora's industry it is *"always about staying relevant to what people are considering"* and being able to accurately detect these trends would be very relevant for Pandora to *"stay ahead of the game"* (Thomsen, 2019). Another example was given by Dülfer (2019) when he refers to the trend of cronuts[2], which went viral in Manhattan a while ago. Through incorporating the described BDA into companies' NPD processes, a trend like the cronut could be identified by companies at a time when they are still *"very locally and small at the start"* and would be incredibly important, especially for idea generation (Dülfer, 2019).

Barfort (2019) and Blom (2019) both suggested the second use case for this theme, whereby a company can use BDA to better understand people's sentiment towards certain trends. By using sentiment analysis on community data companies can analyse if particular segments of people have positive or negative opinions about a given topic. Barfort (2019) provided the example of an initiative by Nike whereby they *"monitor how people discuss Nike only in Los Angeles on I think Instagram and maybe Twitter, and then they actually use that kind of information to build new products that are only sold in Los Angeles."* This example uses real-time sentiment analysis on social media and then the insights generated are given to designers to create the next release of products for that store. Whilst in this example Barfort (2019) is highlighting the potential of such *"social listening"* to identify people's sentiment, he also stressed that the insights derived from this type of analysis can be very general and lack the necessary granularity for designing products. He thus does not call into question the general applicability of this analysis for NPD, but its effectiveness and impact.

[2] A combination of a croissant and a donut.

5.1.1.2 Customer Preference Identification

The *Customer Preference Identification* theme includes various use cases where companies can, by using BDA, better identify and distil the key customer preferences which act as the starting point for their new product ideation. This theme considers customer preferences to include preferences from both existing and potential customers. Seven of our interviewees provided insights as to how BDA could help to identify, understand and distil customer preferences, making it the most discussed theme in the conducted interviews. The use cases in this theme are all similar in that they involve obtaining insights from customer data, however a key distinction between two groupings is in the approach to obtaining those insights. One group can be identified as following an 'outside-in' approach whereby a company will analyse the market using external data and try to extract relevant preference trends that are applicable to its customer base. In contrast, by following the 'inside-out' approach a company analyses its existing customer base using private data to distil the key preferences that drive its current sales and then use that as the basis for its ideation, which could also be aimed to attract new customers. Whilst many of the themes identified through our interviews share some commonalities in the type of data required or the analytics methods used, this theme is notable for containing no such similarities. Within both groupings it is possible to use a mixture of data types, such as internal vs external or structured vs unstructured, and also a number of different analyses ranging from descriptive up to prescriptive. Due to the lack of similarity in the analysis or data, the following will present the use cases independently of each other, starting with the 'outside-in' method followed by the 'inside-out' one (Andersen, 2019; Antille, 2019; Barfort, 2019; Dülfer, 2019; Shah, 2019; Thomsen, 2019). In total six use cases were identified in this theme which are split equally between the two approaches.

<u>Outside-in:</u>

Three use cases could be grouped into the 'outside-in' approach, which were suggested by Antille (2019), Andersen (2019), Barfort (2019) and Dülfer (2019).

Antille (2019) highlighted preference mapping as one use case employed in the outside-in approach, whereby a company can *"identify what are <u>really</u> the properties that a product should have to delight the consumers"*. He adds that this method allows a company to identify *"what products are liked or disliked"* and also why this is the case. In other words, it enables the *"liking drivers"* of particular products to be identified. Furthermore, by using preference mapping to go 'outside-in' it allows a company to *"learn about the different patterns or preferences that exist in your population"* and not just the ones in their customer base (Antille, 2019). Antille (2018) cited the example of using preference mapping for coffee, whereby consumer preferences were mapped according to liking patterns, resulting in three main clusters of coffee preferences termed *Strong, Delicate* and *Indulgent*. From these preferences existing coffee products in the portfolio could be compared to assess how closely they matched these preferences and identify whether a new product needed to be developed. The output from the preference mapping process would act as *"a brief to the development team"* that outlines which properties are valued by consumers and so should be targeted during development (Antille, 2019). Whilst Antille (2019) provided the use case of coffee he stated that this technique could be applied in a number of areas, including skin care for instance, and the only real limit to its application was where the target consumers were not able to provide valid feedback about the product, such as with products for babies for example. It is important to mention that Antille (2019) explicitly said that this technique is only used on product characteristics, thus it does not cover any issues around the price or the packaging for example making it only relevant for the Ideation stage and not the Product Launch stage.

Andersen (2019) provided a second use case where player reviews of games were analysed to extract *"useful insights into what [...] people like and do not like about the games"*. Through text analysis and natural language processing it was possible to extract information from the player reviews which contained unstructured free text. This data was then analysed via machine learning algorithms to identify patterns in the reviews which would highlight preferences for games in the whole market.

Dülfer (2019) and Barfort (2019) both contributed the third use case for the 'outside-in' approach, which revolved around a fashion start-up that was able to identify *"how should the next t-shirt look like"* through social media analytics. Through the use of image recognition on the *"massive amounts of Instagram pictures"* it would be able to cluster commonalities in the images posted, and thus identify patterns which could represent a future trend (Barfort, 2019; Dülfer, 2019). However, unlike the use cases above, the output from this technology would not only predict what were the next fashion trends, it would also *"create the full design"* that was ready to be tested in the Product Development phase (Dülfer, 2019). Despite the fact that this use case transcends the boundary between Ideation and Product Development, it has been included in the idea generation and selection stage as the major benefit to companies is thought to be here.

Inside-out:

Three use cases falling into the 'inside-out' approach were suggested by Thomsen (2019), Shah (2019), and Antille (2019).

First, Thomsen (2019) stated that there can be *"hundreds of factors that impact how a product performs"*, which is why the driving motivator to use BDA is to overcome the complexity inherent in the process of distilling the product attributes that predict success (Thomsen, 2019). Technologies linked to BDA have much greater processing power and so are able to process these potential parameters whilst with a software like Excel it is *"very difficult to see what is the actual driving power, or parameter"* (Thomsen, 2019). Furthermore, Thomsen (2019) highlighted the benefit of being able to *"incorporate other sources of input that are not necessarily just historical sales data"* through the use of BDA, which would allow them to better understand what their target group is interested in. The example cited by Thomsen (2019) showcases what attributes of jewellery charms drive their sales performance, which is very difficult due to the number of factors involved that go beyond the product itself. Thomsen (2019) exemplified*: "Is it because it was silver? Is it because it was a horse? Is it because they didn't put it in the window? Is it because a competitor had something similar? Is it because it was raining all of July and no one went downtown to go shopping?"*

Second, Shah (2019) provided a use case where it was possible to identify different groups of customer preferences by analysing usage data of the product, and then these groups could subsequently be targeted as a separate segment. The example cited by Shah (2019) presents the wearable technology company Fitbit, which initially launched a product at a high price point that came with a number of features included. However, by analysing self-quantification data, derived from the usage of the product, they could see which of their customers were *"actually only looking at sleep, which are the customers that are only looking at the how many steps they are taking, which are the customers only looking at your heart"* (Shah, 2019). Based on these insights Fitbit was then able to launch different product versions with diverse functions and at different price points, targeted at each of the sub-segments. *"Then suddenly from a $200 one then you launch another one which is just $50 but for the market segment which is only interested in measuring the steps and then you've launched another thing for $70 for the segment who are interested in measuring the heart rhythm"* (Shah, 2019). Therefore, through BDA, a company is able to analyse one particular customer segment, typically a higher priced segment with many features, and then through very detailed analysis identify other market segments which provide the ideas for further product development (Shah, 2019).

Finally, Antille (2019) offered the third use case in the inside-out approach as to how, through BDA, a company can better interact with their customers and gain customer feedback about their preferences directly, and in real-time. Antille's (2019) example is related to coffee and states that in Japan they have released coffee machines that come with coffee cups which contain radio chips that act as a sensor. Such a chip allows the company to match coffee consumption patterns and preferences with data on the consumer (Antille, 2019). Furthermore, by connecting this sensor with a user interface, such as an app, this would allow the consumer to provide direct feedback directly on the quality and taste of the coffee. Therefore, this system would allow a company to know the type of coffee that was consumed, by whom, at what time of the day and also provide direct feedback. Antille (2019) offered two possibilities for this self-quantification data, first that the coffee machine could alter the preparation of the coffee to better suit the preferences of the individual – effectively a personalisation of the coffee. Additionally, Antille (2019) suggested that this data could then be used in the NPD process as the basis for idea generation.

5.1.1.3 Input Scanning

The ability of BDA to help companies to constantly monitor the market for emerging innovations is the central idea of the *Input Scanning* theme. The emerging innovations can be, for example, new materials that can enable the company to further improve existing products or to create completely new products (Blom, 2019; Dülfer, 2019). Given that the nature of this theme requires scanning of 'related fields', it follows that the use of BDA in this theme relies on data which is external to the company. As specific sources, Dülfer (2019) mentioned structured data from databases, such as patent databases or publication databases, that can be screened by companies. Whilst no specific type of analysis was given for this theme, Blom (2019) stated that artificial intelligence can be used to scan the market for input on an ongoing basis. The potential of BDA to identify new possible input factors for NPD was emphasized through two use cases (Blom, 2019; Dülfer, 2019). Whilst the first use case highlights how *Input Scanning* could be used to obtain a cost reduction the second shows how it could enable a company to come up with completely new products.

As a first use case, Dülfer (2019) mentioned that this type of BDA is used by companies working with chemical molecules. He cites the case of a refrigerator manufacturer that was screening databases of publications to look for new molecules that could be used within their products. The result was that the company was able to identify an alternative to their current cooling substances that was considerably cheaper, whilst still fulfilling the same requirements.

Blom (2019) demonstrated in a second use case how this approach is also leveraged by General Motors (GM) in the automotive industry to identify new materials. He outlines how GM uses artificial intelligence to continually scan the market for new relevant innovations, which then inform the company's product development efforts. Blom (2019) highlighted that this could allow GM to spot new developments in the material industry.

Against the backdrop of this theme, Dülfer (2019) pointed out that one has to make a differentiation between whether companies look for different components or materials, as in the above stated examples, or whether they just generally want to screen for new technologies. While the latter might also be interesting for firms to identify, Dülfer (2019) emphasized that if you do not really know what specifically you are looking for you also do not *"have any kind of keywords for that and that makes it very difficult in big data analytics to run these things".*

5.1.1.4 Quality Improvement

Two of our interview partners, Shah (2019) and Ahlbrand (2019), mentioned a theme that describes the use of BDA to identify the improvement potential of components used in existing products which can lead to marginal enhancements when these components are used for NPD. The data used in the *Quality Improvement* theme can be both internal and external data and can be classified as either private data in the form of, for instance, machine performance data or self-quantification data. The type of analysis can be described as being either descriptive or diagnostic (Ahlbrand, 2019; Shah, 2019).

The single use case for this theme was mentioned by Ahlbrand (2019), who explained that when it comes to the development of new machines, CLAAS strives to obtain the right to gather private usage data from machines currently running in the field. This data can then be evaluated to obtain information about the performance of the machines and, above all, individual components in order to then improve them. He outlines that *"everyone has different interests in the data"* and that for instance *"the engine department just tries to evaluate the engine data"*. Insights derived from such an analysis subsequently flow into the NPD process of new machines. Similarly, Shah (2019) also highlighted the importance of receiving data from products that are connected in the field as thereby it is possible to, for instance, identify components that are always failing after a short time and improve them so that a new version of that product can be launched. He mentioned that the reason why one should start analysing parts in current products is because you can directly learn from the data (Shah, 2019). With regards to such an analysis, which often involves complex technical data, Ahlbrand (2019) stressed the importance that data scientists collaborate with colleagues that are knowledgeable about the respective field. He mentions that this is essential, because the result of the analysis *"has to be checked for plausibility first"* to assess if the derived measures are realistic or if the analysis was somehow faulty.

5.1.2 Big Data Analytics in the Product Development Stage

As outlined in the literature review, the Product Development stage encompasses the development of initial prototypes, testing these prototypes and then utilizing the feedback to arrive at the final physical product (N. Bharadwaj, 2018). In light of this definition, 11 out of the 12 interview partners emphasized the benefits of incorporating BDA into the Product Development stage and nine were even able to provide a specific use case on how this is already being done in practice. Only one interview partner said explicitly that it is not possible to see how BDA can help in the Product Development stage (Table 4).

In general, Dülfer (2019) said that based on his experience he thinks that BDA can help companies in the Product Development stage, but that the "*advancements are not that far compared to the beginning and the launch phase*." Many respondents stated that BDA is especially helpful in light of the increased use of the lean approach in companies' Product Development stage. The lean approach advocates to build minimum viable products and then test these prototypes early in the market instead of waiting until a final product has been developed before testing it (Andersen, 2019; Shah, 2019; Vas, 2019; Barfort, 2019). Shah (2019) stressed that data and analytics greatly improve such a lean approach as BD provides much better information that can be processed more quickly and accurately by analytics. This allows for the risk of failure to be significantly reduced with the same financial investment in risk mitigation compared to not using BDA, thus accelerating the overall lean approach. With regards to the lean approach, Barfort (2019) stated that he sees BDA as a "*way to kind of scale prototyping"*, as prototypes can be tested early on and through BDA companies will obtain feedback on a scale where it will be trustworthy quite quickly.

Besides these general comments, several interview partners were also highlighting the potential of BDA for certain activities in the development stage such as for design and testing. With respect to the design of products, the opinions as to what extent BDA can support this process were quite diverse. At one end, Vas (2019) stated that product design can *"100%"* be influenced by BDA and that in the future machines will be able to build better products based on data than product designers. At the other end, Barfort (2019) said that he sees very limited instances in which designers can be substituted by machines and see the two as being complementary.

With respect to product testing, Barfort (2019) mentioned that by using BDA companies can test their products with larger quantities of customers and thereby get a higher generalisability in their results.

He points out that traditional methods of testing, where you manually observe how potential customers interact with the product are unlikely to this level of generalisability. This is in line with Shah (2019), who exemplified that BDA, specifically machine learning, enable companies to reach the same quality in tests after only six weeks, while traditional testing could take six months. This would support the decision-making process around the yes or no decisions on whether a product is going to succeed or not (Vas, 2019). However, Barfort (2019) also stated that it still represents a challenge to test products with BDA, except if they have some form of software component as many tests still rely on some form of self-reporting. Through the increase of IoT devices it will, however, be much easier to test numerous products on a large scale over the coming years.

From the interviews four themes could be identified that are influencing at least one of the three NPD performance criteria. These four themes are: *New Product Verification, New Product Exploration, Customer Preference Matching and Performance Optimisation.* As with the Ideation stage, these themes will be explained in detail and use cases outlined, and they will be presented in the sequence they would occur in the second stage of the NPD process. Hence, the subsection will start with the two themes that focus on design and development, *Customer Preference Matching* and *Design Optimisation*, followed by the two themes that focus on testing, *New Product Verification* and *New Product Exploration*.

5.1.2.1 Customer Preference Matching

The unifying idea behind the *Customer Preference Matching* theme is that once customer preferences are identified, companies can leverage BDA to create products that, to a great extent, satisfy these preferences. For this theme, customer preferences need to be either identified by the company, as illustrated in the *Customer Preference Identification* theme, or stated by the customers themselves. Based on that, data can be used to model product properties with the intention to match the identified customer preferences. BDA is then used to compare the identified preferences with the proposed product attributes, which are thereafter automatically refined based on the results of the analysis. This process would continue in iterations until a product is developed, which largely meets the customer preferences (Antille, 2019a; Hede, 2019). This theme is illustrated by two use cases which were provided by Antille (2019) and Hede (2019).The two use cases can be distinguished according to the complexity of the data analytics involved, with the latter use case requiring more advanced techniques for analysis such as machine learning.

The first use case is provided by Antille (2019b) who explained that an ANOVA[3] based data analytics approach is used at Nestlé for genuinely new product innovations but also incremental product improvements across a variety of product categories and level of design complexity. Antille (2019a) explained that simply knowing the customer preferences does not automatically mean that product development knows how to develop the product. He exemplified that for a new coffee product, it is of course beneficial to know that *"the product needs to have this thickness, this sweetness and this acidity, but then how do you make it?"* (Antille, 2019a). This is where data analytics comes into play (Antille, 2019a). Based on preference maps, which were explained in subsection 5.1.1.2, Nestlé is using statistical modelling to process structured private data about how, in the case of coffee, different coffee types, dosages and machine parameters change the sensory properties of the resulting product. To emphasize the advantage of data analytics here, Antille (2019b) stated that in some cases only 32 experiments need to be performed due to the use of BDA instead of 36,864 which would otherwise be necessary.

As second use case is outlined by Hede (2019) who explained that generative models, which are a form of unsupervised learning, can be used to match a certain set of preferences. Therefore, these models have to be trained using data about similar products like the one that is in development. By providing the model with huge amounts of unstructured data it can analyse and thereby learn the structure of the data. Hede (2019) stated that once the model understood the structure it could then design an example of that product with the necessary characteristics. While the design of the product itself would no longer qualify as BDA, the design is performed on the basis of the specifications that could be derived from the BDA. To illustrate this method, Hede (2019) described the case of Adobe, which is not a product company but can exemplify the potential of generative models. Adobe trained the model with millions of different fonts so that now customers "*can tell it like I want a san serif, monospace, that looks good in large print and then it just generates it for you"* (Hede, 2019). Hede (2019) states that this font would be completely new and unique and further envisions that similar things are possible for other products, for example cars. With regards to personalisation Hede (2019) stated that "*instead of having 20 different models you can have infinite number of models and just generate variations based on customer requests*". However, he conceded that it might not yet be possible to fully design such complex products due to the current technological development and that initially it will probably be half generated, and the other half finished by a human who brings it all together.

[3] Analysis of Variance

5.1.2.2 Performance Optimization

Another theme, *Performance Optimisation*, focuses on the improvement of product designs with respect to certain performance criteria. In contrast to the theme presented above, *Performance Optimisation* is not oriented towards meeting customer preferences, rather it is only concerned with designing the product to maximise performance along a given dimension. As the product design is being evaluated against fixed performance criteria the data for this analysis is typically of a structured nature (Hede, 2019). There was only use case in this theme, which was provided by Hede (2019).

Hede (2019) stated that the method of generative design can be employed to achieve this goal. Generative design is a design process whereby design goals are inputted along with parameters for performance, spatial requirements, materials and constraints and then all potential design options available for the inputted data will be explored. With each further iteration it will learn automatically from the structured data and can further improve the design (Hede, 2019). A concrete example that was provided by Hede (2019) is the design process for propellers. Generative design allows for *"really good physics simulations"* and it can *"generate billions of designs"* and thereby is capable of coming up with radically innovative propeller designs with a better performance. At the same time, it must be noted that Hede (2019) cautioned against the unbridled success of generative design because, despite generating many propeller designs, none were possible to produce as the simulation wasn't accurate enough. But in general, it might be an approach to come up with designs that are *"radically different from what we know today"* (Hede, 2019).

5.1.2.3 New Product Exploration

In the *New Product Exploration* theme, BDA is used to broadly explore how customers think about and interact with product prototypes and thereby improvement potentials can be identified. To achieve this qualitative and unstructured data is used to identify how products can be further improved, without having a clear performance or design objective in mind. Two use cases were identified that fit with this theme (Barfort, 2019; Shah, 2019). Whilst with the first use case BDA provides the necessary insights to be able to design a new product to a designer, the second use case goes further and suggests that the entire process can be done without human intervention.

Barfort (2019) provided the first use case of testing a newly developed chair to identify how people interact with the product. By setting up these chairs in a live environment, for instance in train stations, airports or other public places, it would be possible to derive self-quantification data on how people sit in the chair, how often they would move around and change their position in the chair through sensors or cameras. Barfort (2019) noted that this is information companies would otherwise likely be unable to gather, as "*all of us can see our friends and family sit in chairs for a few hours, but we can't do that for thousands of people for many days*". Thereby, companies would receive scalable data on how customers interact with their chair and how they use it in different ways. This highly unstructured and qualitative data can subsequently then be analysed by algorithms using among others video analytics. Barfort (2019) outlined that the first thing these algorithms would do is "*basically just to map some sort of typology of how people are using the chair. What are the patterns of ways that people are moving around in the chair?"* However, according to Barfort (2019), the algorithms would not be used to give specific recommendations about how the chair could be improved, as this is something "*that humans would be much better at than the machines."* Instead, these algorithms would be more descriptive in nature and identify patterns and differences in people's behaviour in how they interact with the chair. This can already fill gaps in designers' knowledge, that are otherwise very hard to fill and can provide valuable insights for the next design iteration.

A second use case that fits into this theme was mentioned by Shah (2019). He outlined that there are beer companies starting to experiment with artificial intelligence (AI) in their testing of new products. Companies would actually serve first prototypes of the beer to customers who would then have the chance to state via an app how they liked the beer along with demographic information about themselves such as the age group. All this information would then be used by an algorithm to identify via prescriptive analytics how the next iteration of the beer should be like. This whole process of testing and refining the product based on the insights is done without human intervention (Shah, 2019).

5.1.2.4 New Product Verification

Besides *New Product Exploration* explained above, another theme was found related to product testing. While *New Product Exploration* makes use of qualitative and unstructured data to find out how products can be further improved without a clear objective, *New Product Verification* leverages more quantitative and structured data to analyse if a product meets certain pre-defined requirements. If a product meets the pre-defined threshold then it will pass through the Product Development stage gate and move on to Product Launch. The type of data used for this theme is of a quantitative and structured nature. Three use cases were identified that fall into this theme, based on insights provided by Ahlbrand (2019), Dülfer (2019), Shah (2019) and Thomsen (2019). The three use cases are distinguished according to the specific type of verification that is being undertaken, namely, technical, regulatory or commercial.

With regards to technical assessment of the products, Ahlbrand (2019) mentioned that BDA is *"increasingly becoming a very, very important feature or tool in testing"*. He outlines that for testing of newly developed machines his company uses telemetry data transmission *"to generate the evaluation at the same time, so that the colleagues from the testing department can directly view the measurement results the following day"*. This testing approach is drastically different from earlier testing approaches, where for instance data loggers were put on the machines and it would take weeks until evaluations of the private usage data could be carried out. The big advantage of telemetry data is that based on the timely evaluations, tests can be refined, and missing measures can be obtained quickly. In contrast, before it was often the case that at the point where missing measurements were identified, the machines would have been removed already from the test sites (Ahlbrand, 2019). However, to analyse the data properly and in a time-efficient manner, dashboards have to be built before the telemetry data arrives in order to be able to analyse the correct data and visualize it in an appropriate way. Dülfer (2019) concurred with Ahlbrand (2019) and added that in the automotive industry, engines that are supposed to run at least 200,000 kilometres are not tested for this distance. Instead, companies are able to draw conclusions about the performance of the engine by statistically modelling various engine data points. According to Dülfer (2019) this type of analytics would fall into the category of predictive analytics and could also be used to analyse if the car or components of it meet certain thresholds as defined by regulation.

Besides the technical testing, regulatory testing seems to be a second concrete use case, which was mentioned by Shah (2019) in the context of pharmaceuticals. The pharmaceutical industry contains a particularly extensive compliance process and reading through the corresponding compliance documents is a time intensive task. Shah (2019) highlights that through BDA these documents, which represent public data, can be screened and based on an algorithm the probability that a certain pharmaceutical product and its characteristics would violate a certain regulation could be assessed. Thereby, the readiness of the product for the market can be verified (Shah, 2019).

In addition to technical and regulatory verification, commercial verification is a third concrete use case, which was mentioned by Thomsen (2019). Together with a partner company called FirstInsight, which specialises in digital product testing via predictive analysis (FirstInsight, 2019), Pandora is testing the newly developed jewellery products online. This is being done for nearly all newly developed products, except those which are deemed to be very basic and for which the uncertainty about their market potential is not that substantial (Thomsen, 2019). Together with FirstInsight, Pandora is designing very simplistic surveys that do not focus on qualitative feedback, but rather on structured quantitative feedback of how customers like certain products. The high simplicity of the surveys allows Pandora to obtain massive amounts of feedback, which is then analysed via BDA. The feedback obtained via these surveys represents one of the central decision criteria as to whether to move the products on to the Product Launch stage (Thomsen, 2019).

5.1.3 Big Data Analytics in the Product Launch Stage

As defined in the literature review, the Product Launch stage encompasses launching the product in the market and ensuring that it is a commercial success (R. Adams et al., 2006). In light of this definition, 10 out of 12 respondents emphasized the benefits of incorporating BDA into the Product Launch stage, and six were also able to provide a specific use case. The other two interview partners did not deny the benefits of BDA for the Product Launch stage, but due to their position within their companies they lack the insights into these more marketing-related activities and therefore did not want to commit themselves to a clear statement (Table 4).

The potential of using BDA for the Product Launch was especially highlighted by Dülfer (2019), who stated that the technologies that can be used within the activities in this stage are so advanced that they are already used on a broader scale. Dülfer (2019) and Barfort (2019) emphasized that leading media agencies show what is already possible by using BD, as all their clients' marketing mix and channel decisions about how to reach their customers in the best way is already being done via BD and being monitored on a continuous basis. Dülfer (2019) put into perspective that this *"advancement will increase further in the next months and years to come."* Barfort (2019) explained the high applicability in this stage by stating that activities within the Product Launch are quite straightforward to quantify and therefore incorporating BDA in these activities is similarly straightforward. He underlined that "*the reason analytics works very well for these kinds of things is that for such activities it is very easy to do controlled randomized experiments, especially online*". While A/B testing for product design is a suboptimal and expensive approach due to the infinite number of product characteristics, A/B testing in Product Launch activities is much cheaper and faster (Barfort, 2019). Hede (2019) outlined that in particular AI *"could do a lot of sort of super charging"* of A/B testing in the Product Launch stage and can automatically learn about what are the best launch decisions for instance in relation to product pricing. While it is already implied here that these activities take place without human intervention, Dülfer (2019) explicitly stressed that such activities will be done in the future most likely autonomously. He states, that *"probably there will be even like a war of robots which will actually make or break the launch of a new product I could imagine."*

Based on the interviews three themes could be identified that are influencing at least one of the three NPD performance criteria. These three themes are *Promotion Tailoring*, *Price Determination* and *Launch Calibration*. The subsection will begin with the two themes that focus on activities to figure out how a product shall be launched, namely *Promotion Tailoring* and *Price Determination*, before presenting the third theme on *Launch Calibration*.

5.1.3.1 Promotion Tailoring

The *Promotion Tailoring* theme is concerned with how companies can use BDA to tailor a company's promotion for a newly launched product either to specific customer segments or even to individual customers (Ahlbrand, 2019; Barfort, 2019; Dülfer, 2019). The common premise of these use cases is that a company's promotional activities are more specifically targeted to customer segments or individual customers. However, the use cases demonstrate that this can broadly be achieved either by tailoring the content of the promotional material or by more carefully selecting the customers so that the promotional activities are a better fit. Therefore, this leads to divergent uses of data and analytics techniques as tailoring the promotional material must happen in real-time using machine learning algorithms, whilst selecting the customers can occur on a rolling basis (Ahlbrand, 2019; Barfort, 2019; Dülfer, 2019). Ahlbrand (2019), Barfort (2019) and Dülfer (2019) all provided insights into three use cases that are within the scope of this theme.

Dülfer (2019) noted that companies are now able to customise products from a production perspective, creating *"Lot 1 products"*[4], but increasingly companies are becoming more intelligent in optimising the *"sales mix and packaging according to the [customer] preferences"*. Barfort (2019) echoed this sentiment and added that this occurs due to machine learning algorithms which are able to adjust the marketing material in real-time *"depending on how a consumer initially reacts"*. Such advances in tailoring of promotional activities have been enabled by new technologies, such as eye-tracking technology, which allow companies to monitor how consumers interact with their websites and promotion materials (Barfort, 2019). Dülfer (2019) provided first use case concerning an advertisement for a basic product which, when marketed via a *"high fashion or super luxury"* channel the marketing material around the basic product would be enriched so that *"people feel attracted to it"*. However, when advertised on a gaming page the consumer would see *"completely different ads with the same basic product but super stripped"*. Barfort (2019) provided the second use vase which was a slight variation on the previous example as he suggests that company websites could be built with this tailoring in mind, whereby machine learning algorithms *"dynamically generate the website"*. Therefore, a website could be fundamentally different depending on how a consumer interacts with

[4] This is a term used to refer to very small quantities of goods manufactured in a single production run.

the page. Whilst the variation in tailoring is somewhat limited now, Dülfer (2019) sees this trend continuing and believes that in the future the tailoring will be even more dramatic whereby the *"people that you see on the ad will change"*.

Ahlbrand (2019) alluded to a third use case within this theme, whereby companies are able to combine different data sources to be better able to target which customers to launch new products to. Unlike the use case above where the marketing material is tailored to specific segments here the focus is on identifying the precise customers who are most in need of a new product and so should be contacted first. Essentially, Ahlbrand (2019) suggested that by combining customer data from a CRM system with telemetry data it is possible to identify which exact customers are most in need of a new product. The CRM system provides the qualitative and transactional history of customers, inputted by the sales team, and the telemetry data is gathered directly from the machines about their usage and so provides insights into which machines need replacing and when. The resulting insights allow firms to optimise their sales mix decisions when launching new products.

5.1.3.2 Price Determination

This theme is concerned with setting the optimal price for a newly developed product. Setting the optimal price is based on a number of factors and can vary for each product and for each customer segment, however this theme is concerned with setting the optimal price based on the required objectives and parameters set (Barfort, 2019; Hede, 2019; Thomsen, 2019). A common component of this theme is that the price setting is conducted based on customer data, namely customer feedback. However, a key difference which separates the two use cases is that this customer data can be obtained actively, by the customer providing a suggested price for a product, or passively, by deciding to buy a product at a given price point. Implicit in this distinction is that when the price setting is based on a customer decision this involves analysing real-time data, whilst when the customer is providing the price, the analysis can be done in stages (Barfort, 2019, Thomsen, 2019). In both cases the data is of a structured nature and quantitative as either a customer would recommend a price figure or there would be a binary data point based on the customer decision (Barfort, 2019; Hede, 2019; Thomsen, 2019). Barfort (2019), Hede (2019) and Thomsen (2019) elucidated the two use cases that fall within this theme.

In the first use case, provided by Thomsen (2019), potential customers are shown images of the product online without any price points and then customers are asked *"how much they like it and what they're willing to pay for it"*. The subsequent analysis of this customer data would consist of a series of machine learning algorithms which are applied to derive a price sensitivity curve and then identify the most accurate price point for the given product (Jezerc, 2018; Thomsen, 2019). The main benefit derived from this digital method of price setting, according to Thomsen (2019), is that *"it's just super simple"* in comparison to more traditional survey methods. Consequently, the simplicity means that you get more answers from customers and thus the validity of the price sensitivity curve is increased, and a more accurate price can be set for the product.

The second use case, proposed by Barfort (2019) and Hede (2019), differs from the use case above as it provides products with prices to actual customers in real-time and then based on whether the customer decides to proceed the price is adjusted. In a static one-off setting this process can be achieved using a standard algorithm, providing it is clear what are the parameters and the overall optimisation objective of the algorithm what the algorithm. However, Hede (2019) noted that the type of analysis required to perform this in real-time is called reinforcement learning, which is where AI would optimize the price based on a goal but *"that goal might not be predefined in data so it might be something that the AI learns as we go along"*. Essentially, this means that the AI would give a customer a price and see if they purchase the product and then based on this decision it would *"learn by itself what is a good price of the given product"* (Hede, 2019). Whereas the use case proposed by Thomsen (2019) can only be applied on a segment level, this second use case could enable it to even calculate a suitable price for a specific customer (Hede, 2019).

However, Barfort (2019) highlighted the challenge of pricing a new product this way. The issue is that by *"taking an atomistic approach"* to pricing, it is possible to forget to factor in the signalling effect which is created when setting a price. Barfort (2019) cited the example of Apple being able to price their smartwatch at a lower price to generate more revenue but refusing to do so due to the potential to *"put downward pressure on Apple iPhones"* and subsequently alter people's perception of the Apple brand. Therefore, Barfort (2019) argued that it is not as simple as it may appear and that it is crucial to know *"exactly what goes into the algorithm"* to avoid any unexpected *"downstream consequences for your entire product portfolio."*

5.1.3.3 Launch Calibration

The central concept around this theme is how BDA can be used in pilot launches to ascertain if the choices made for the planned product launch are optimal. Whereas the other two themes within the Product Launch phase cover a particular set of decisions (i.e. product promotion or pricing) this theme provides a holistic view and represents the final activity before the product is fully launched (Barfort, 2019). Consequently, this theme cannot be characterised by any one type of data or analytics approach. Barfort (2019) provided insights into the use case which forms the basis for this theme.

Barfort (2019) argued that companies are now able to launch products in specific environments and *"track customer behaviour"*. The notable element of pilot launches is that they assess a wide variety of elements around the product holistically, therefore a variety of different data sources is required to be able to get the whole picture of the customer's interaction with the new product. By only using internal data, such as consumption data or surveys, or external data, such as social media posts or customer reviews, it is not possible to get the holistic picture necessary to judge a pilot launch (Barfort, 2019; Dülfer, 2019). Dülfer (2019) highlighted the importance of data to get this holistic picture when he said *"there are some things which you can get out of the data, which you simply would not get out if you ask people"* even potentially things that *"they are completely unconscious about"*. Barfort (2019) provided the example of launching a new type of milk and by utilising BDA a company could *"randomize, take 10-15 supermarkets in Copenhagen, offer these kinds of milk on the shelf and then they could basically track customer behaviour right there in the store"*. The variety of data is evident as the company could get access to the consumption data and complement that with unstructured data from cameras and sensors in the stores (Barfort, 2019). Whilst Barfort (2019) acknowledged that getting the variety of data necessary to judge pilot launches is much easier when the product contains a software component, he says that in a few years this will also be much easier for all products due to the proliferation of smart home devices.

5.1.4 Identification of Further Use Cases

In the subsections above there are 26 individual use cases which are grouped into 11 overarching themes and spread across all three stages of the NPD process. While this selection of themes provides a useful overview of how companies can integrate BDA into their NPD processes, underpinned by concrete use cases, this typology does not claim to be exhaustive. The authors acknowledge that there might be use cases that would not fall into one of the presented themes but that would create new themes. In light of this, the authors consider it worthwhile including the insights, provided by Hede (2019), into the criteria for identifying a good use case for using BDA in the NPD process. Hede (2019) stated:

> *"You need to figure out somewhere in an organization that if someone knew something that they don't know today, they could do something differently. […] We would like to give people a crystal ball, and be like, 'Where in your daily work - let's imagine this crystal ball worked, let's imagine it could actually predict the future - where in your daily work would you poke it and be like, ask it something, and when would it make a difference?'*

It is in these circumstances, asserted Hede (2019), that BDA can show its true potential for influencing the NPD process and consequently NPD performance. However, Hede (2019) cautioned against this being an easy process as in reality *"it is surprisingly hard for people, when you really put them on the spot, to be like I would like to know that".*

The results above presented the means by which BDA can be incorporated into the NPD process. The next section will build upon the results presented above to demonstrate how incorporating BDA can influence the three NPD performance criteria, namely cost of development, speed-to-market and product-market fit.

5.2 Influence of Big Data Analytics on New Product Development Performance

After having presented the means by which BDA can be integrated into the NPD process, the following subsection will present the results as to how incorporating BDA in the NPD process can influence NPD performance according to specified performance criteria. The cost of development, representing the input factor of the NPD process, will be presented first. The results related to the speed-to-market, which represents the NPD process criteria, will follow afterwards. The final subsection is dedicated to the output criteria, namely product-market fit.

5.2.1 Cost of Development

With respect to the importance of the cost of development in NPD, Barfort (2019) outlined that even though the cost of development "*enters into your kind of whole calculus of figuring out what kinds of products to even consider innovating"* he stated that once companies decide on certain innovation projects, *"costs usually take a backseat"* within the NPD process. He highlighted that it is much more important to get the product right and that only after products are fully developed companies try to reduce costs of producing those products.

Whilst the importance of the cost of development was not stressed through the interviews, six out of seven affirmed that BDA could have a positive influence on the cost of development criteria. Furthermore, different levers as to how the cost of development can be reduced by incorporating BDA in the NPD process were mentioned in the interviews. First, activities can be shortened and thereby the costs arising from those activities are reduced (Anjos, 2019). Anjos (2019) stated that she sees the potential of reducing development costs by shortening activities mainly in the Ideation stage, specifically when gathering information. This is in line with Thomsen (2019), who stated that huge surveys to derive customer insights at the beginning of the NPD process are very expensive. Therefore, by using BDA to enhance the amount of insights that are already available about a company's customers beforehand can significantly reduce the amount that has to be spent on market research (Thomsen, 2019). The subsection on speed-to-market will explain in more detail how BDA can help streamline certain activities in the NPD process.

Besides cost reductions through shortening certain activities, BDA can also enable companies to substitute traditional activities in the NPD process with cheaper BDA-based approaches. In this context, the use case about testing with telemetry data at CLAAS, which is part of the *New Product*

Verification theme, represents a good example. Ahlbrand (2019) outlined that through this new testing approach, far less measurement technology has to be installed on-site. He explains the significance of this change as follows:

> *"These studies or validation measures, they are usually carried out in Australia or elsewhere. Until you have the measurement technology there and set it up and until it all works, there are also high costs involved and if we can do all this with telemetry data, we can reduce costs considerably."*

Besides shortening and substituting activities, costs can also be reduced through eliminating wasteful actions. According to Shah (2019), costs of development are generally not driven by material input but instead by the amount of experiments you conduct to figure out the right set of features and the right materials for your products. Shah (2019) stated that you can apply BDA to be much clearer about what you want and can do. As the use case about identifying compliance guidelines, mentioned in the *New Product Verification* theme, showed, deploying BDA can help you identify the right actions to do, and will thereby eliminate wasteful actions. This can reduce your costs of development, or instead it could also "*give you a better return on investment because for the same cost you can actually deliver more to the customers*" (Shah, 2019).

While these statements cover the Ideation and Product Development stage, Barfort (2019) stated that in general he believes that BDA can also play an important role in the Product Launch stage and can reduce the costs in this stage. According to him, it is "*very clear that [incorporating BDA] puts downward pressure on marketing agencies"*, which are nowadays managing the launch and marketing activities for many big corporations (Dülfer, 2019). Despite this affirmative statement, no other assertions on whether BDA can help reduce the costs-of-development in the Product Launch stage were provided.

Whereas the levers by which costs of development can be reduced are clear, the improvements that are generated are hard to measure for two reasons. First, the interviewees mentioned that with costs of development it is difficult to assess the improvements through BDA as it is difficult to measure. Oftentimes it can be assessed that through BDA the time and quality of, for instance, gathering insights in Ideation is improved, but how that affects the costs of development is hard to say (Anjos, 2019). Anjos (2019) stated that at Coloplast it is not possible to put a *"figure on that, at least not today"*.

Second, there are different time dimensions which can be used for cost improvements, which makes it hard to assess if BDA can lead to cost-savings. According to Shah (2019) it depends if you look at the set-up costs or the operational costs. He states:

> *"So, it depends if you just look at the investment cost, of course you have got to invest a bit more. But if you look at the overall cost in say a 5-year timeframe, then it will be lower. But if you look at the first-year costs it will be higher. So that's the tricky part so when it asks about cost, what is the time frame?"*

That is why it is not possible to give a generalisable answer to whether incorporating BDA reduces the cost of development as you always have to consider the timeframe (Shah, 2019).

5.2.2 Speed-to-Market

The next subsection will outline the importance of speed-to-market for product companies and will assess if and how incorporating BDA in the NPD process can streamline the NPD process.

The general importance of higher speed-to-market for product companies was emphasized through the conducted interviews. Shah (2019) highlighted that this is mainly based on a shift in society, as especially young people do not have a long-term perspective and always want the latest products or technology. He outlined: "*My parents had a fridge which is still there, it's been operating for 40 years and if I tell anybody now you have this car for five years, they will say what's wrong with you*". The consequence of this trend is that product companies' speed-to-market becomes extremely critical because customers are moving so fast in terms of their buying patterns, that companies cannot afford to only launch a new product every three years (Shah, 2019). Anjos (2019) also highlighted the importance of speed-to-market, as she said that in her industry first-movers usually have an advantage.

In general, five out of seven interviews confirmed that BDA can help product companies to increase their speed-to-market. In terms of specific activities that can be streamlined, Anjos (2019) stated that especially the time gathering information in the Ideation stage can be reduced from for instance six down to one or two months. Vas (2019) mentioned that in his opinion in the Ideation stage BDA cannot completely substitute the need to speak with customers. However, incorporating BDA in line with the themes *Trend Spotting* and *Customer Preference Identification* can cut down the necessary interaction with customers as there will be not many uncertainties left and that "*will probably save [companies] 50% of the time*". Identifying societal trends quickly, like in the *Cronut* example mentioned earlier, also has a high strategic importance as it allows companies to identify and incorporate these trends into their NPD earlier then their competitors (Dülfer, 2019). Such improvements in terms of reducing the time to derive valuable insights in Ideation, does not, however, necessarily mean that these insights then also ease the progress of these products in the development and launch stages (Anjos, 2019).

However, activities can also be streamlined by including BDA in the Product Development stage. With respect to development, automating product design, a use case discussed in the *Customer Preference Matching* and *Performance Optimisation* themes, can significantly reduce the time of designing the product (Vas, 2019). Vas (2019) emphasized the potential time savings by giving an example of such automated design from the gaming industry. He stated that if companies want to design a game in the city of New York, it would normally take 30-50 days to take pictures of the city and build the game accordingly. Today some gaming companies input the pictures into a BDA model, which analyses the structure of the pictures and then recreates the city in digital form. This can bring down the length of the activity down to three or four days (Vas, 2019). The potential to streamline the product development by using BDA was also emphasized by Antille (2019) on the basis of the previously mentioned use case of Nestlé from the *Customer Preference Matching* theme. Using BDA in such a way to match predefined sensory properties of customers is much faster than to try it with the traditional trial and error approach. Antille (2019) explained that when he said:

> *"If you imagine you are in an innovation unit and you are exploring a new type of product. […] With the models it is like a GPS, you set here is my destination and here is where I am today, show me the fastest way to my destination. If you go with a trial and error, it is like going to a new place that you have never been before with your car with no GPS and you know where you have to go but you have good chance to get lost."*

Besides the development of the product, BDA can also speed up the testing of developed prototypes. In general, Shah (2019) outlined that with the power of BDA the same quality of testing can be achieved in six weeks, as opposed to six months, as machines can process a lot more data much faster. One of the use cases in the *New Product Verification* theme, in which telemetry data is leveraged for testing machine performance, represents a good example. In this context, Ahlbrand (2019) stated that with older testing approaches it "*took weeks until the corresponding evaluation could be carried out*". With the new BDA driven approach, the testing department can view and analyse the measurement results already the following day. Even though the actual time of testing can be shortened drastically, Ahlbrand (2019) also outlined that this BDA driven approach needs certain preparation time upfront, as you have to build dashboards in advance, before you collect the telemetry data and this is "*naturally time-consuming*". This will however only be significant in the beginning, as later on dashboards can be reused.

It was further indicated that not only certain activities in development or testing can be streamlined, but that whole iterations, consisting of testing and subsequent design refinement, can be made faster. An example which illustrates this is the use case from the *New Product Exploration* theme, in which companies are using AI to automatically iterate beer recipes based on the digital feedback of customers (Shah, 2019).

Besides mostly affirmative statements about using BDA in the second stage of NPD, Anjos (2019) said that for Coloplast and their products, she believes that BDA cannot streamline Product Development, as she cannot see how the data can reveal relevant insights in this second stage. However, she stated that in the subsequent Product Launch stage she sees a lot of opportunities to speed up the marketing execution. However, despite this more general statement, the interviews did not reveal any more concrete insights into how BDA can streamline activities in the Product Launch stage.

With regards to the whole NPD process, the interviewees were much more reflective as to whether BDA can significantly increase speed-to-market. Ahlbrand (2019) stated that he sees that there are huge improvement opportunities in terms of the speed-to-market, but that he does not know if CLAAS has already become faster in that regard. This is due to certain learning processes that still take place, as with 10,000 employees a transformation to more data-driven approaches take time. Once BDA is seen as an integral part of NPD, he is sure that it will improve the speed-to-market of products.

Anjos (2019) stated that for Coloplast she does not see how BDA would significantly improve speed-to-market in NPD. She affirmed that through BDA part of the process, especially activities in Ideation and Product Launch, can be streamlined, but that most of the other necessary steps in the process can hardly be influenced. She stated that in the case of Coloplast, the patenting process takes years. Moreover, testing and creating clinical evidence still needs to be done in the same way, which is very "*cumbersome*". This is in line with Thomsen (2019), who also stated that it is not the analysis that is slowing the NPD process down. She states that BDA could help to get analysis done quicker in Product Development on a daily basis, but that in the end it is quality testing and setting up production before the launch that takes up the most time. For these activities it is hard to see how BDA can help and thereby questionable if BDA can increase speed-to-market significantly (Thomsen, 2019).

5.2.3 Product-Market Fit

The ability of BDA to improve the speed-to-market has been explained in the subsection above, however as Anjos (2019) alluded to *"there is a little bit more to a successful product than just make it in the right time"*. Therefore, ensuring that the product itself is also the 'right' product is key, therefore ensuring that there is a strong product-market fit is important. The following subsection will present the results from the interviews conducted as to how BDA can influence the product-market fit.

The importance of ensuring a close product-market fit was also confirmed during the interviews. Thomsen (2019) explained that achieving product-market fit is about *"identifying the need and ensuring that the products we produce actually match that need"* which, in turn, has an impact on a company's top line. It was noted that a focus on ensuring a strong product-market fit is still very important even for a *"cost-conscious company"* (Thomsen, 2019). The fundamental objective of ensuring a strong product-market fit is so that a company doesn't *"do a lot of products that are failed"* (Thomsen 2019). The importance of that was exemplified by Antille (2019) with regards to the food industry, stating that there is an incredible number of products that are launched "*which disappear 6-10 months later just because the consumer research was not done properly*."

Furthermore, six out of seven interviewees affirmed that BDA could have a positive influence on this NPD criteria. Shah (2019) was very positive towards this notion and expressed that product-market fit was the easiest of the three criteria to see how BDA could have a positive influence.

Within the Ideation stage it was clear that BDA could have a strong positive influence on the product-market fit. The ability to identify trends in society was seen as a viable method to improve the product-market fit, which Dülfer (2019) attributed to the fact that it is *"because you know things earlier than others, you detect trends earlier"* which allows firms to improve their NPD on the *"quality side"*. The

Trend Spotting theme above contains some use cases that illustrate how companies are able to achieve these insights into trends which impact product-market fit.

Also falling within the Ideation stage, it was noted that gaining an increased understanding of customer preferences through BDA would allow a firm to increase the fit of a product to a given customer segment. As Dülfer (2019) outlined, *"if you learn more of what your clients want then you can also produce better products, so I think this is key."* Shah (2019) noted the ability to identify customer sub-segments via the analysis of product usage with BDA. The consequence of identifying unique usage patterns is that it reveals customer sub-segments which then leads to the launch of products which are a better fit to the needs of those particular sub-segments (Shah, 2019). Thomsen (2019) added that the process of distilling the key attributes that drive product sales, which can only be achieved through BDA due to the large and complex data sets, is key to improving product-market fit. Distilling these key attributes would increase a firm's understanding of their customers' preferences (Thomsen, 2019). However, this would predominantly improve the product-market fit of the next phase of products launched. In particular, this process would allow a company to identify whether the product characteristics themselves were not a good fit to the market or whether there were other elements affecting sales performance. In turn, this would then *"have an impact on what we [Pandora] would then want to design going forward"* (Thomsen, 2019). The theme of *Customer Preference Identification* is covered in depth above with many use cases highlighting different ways firms could achieve this.

In the Product Development phase, it was noted by Barfort (2019) that a company, by using BDA to analyse behavioural data, could create a product usage typology which could generate insights that lead to an improved-product market fit. The key element which aids the product market fit is that the feedback provided by the behavioural data is at a scale whereby the conclusions can be relatively trustworthy (Barfort, 2019). Barfort (2019) contrasted this BDA-powered method to needing to *"watch thousands of people to interact with prototypes"* to be able to obtain the same validity in the findings. The method by which a company could achieve this was explained in the *New Product Exploration* theme with the Airport Chair use case. Notably, there was no mention of ways in which BDA could be employed to improve the product-market fit during the Product Launch phase.

Outside of the individual stages there were also some interesting insights obtained via the interviews as to how to achieve greater product-market fit. Anjos (2019) stated that *"we never say no to data,* *and never is there enough data"* and that *"we could have even better products if we have more information"*. With these comments Anjos is alluding to the importance of having a large volume of data in NPD to improve product-market fit. In addition, Thomsen (2019) noted that being able to incorporate a variety of data from different sources that are not necessarily just historical sales data would help Pandora to identify what their target group are interested in. This, she argued, would then allow Pandora to *"understand better the needs"* and so create better matched products that were less likely to fail. Anjos (2019) added that as well as having a large volume of data it is very important to obtain real-time data wherever possible. She states it is the *"dream for a company to have customer information pretty much in real time. That's a dream because then you can offer customized products.* (Anjos, 2019). Finally, as alluded to above in the development example utilising behavioural data can also reveal patterns of behaviour that lead to insights which improve the product-market fit (Barfort, 2019). The unifying theme behind these data considerations is that it is only through BDA that these types of data could be analysed. Therefore, the fact that they were provided as desired elements in the NPD process suggests that BDA could play a role in improving product-market fit.

However, some of the interviewees expressed some concern as to whether the benefits outlined above could realistically be obtained. Barfort (2019) expressed concern as to whether the *"granularity"* of the insights obtained from the activities contained within the *Trend Spotting* theme would be very helpful in the NPD process. In essence, Barfort (2019) was suggesting that because these analyses are on a macro-level then the insights obtained are very general and very uncertain. Ahlbrand (2019) added that in the case of CLAAS when it comes to considerations of product-market fit, the *"product manager is still heavily involved"*. Therefore, it is the product manager who identifies the requirements of the market and maps the specification of the product to meet those requirements rather than through any use of BDA. Barfort (2019) also added that the influence of BDA on product-market fit is contingent on the innovativeness of the product. He states that is difficult to use BDA to improve the product-market fit when this involves identifying unmet needs of consumers which is a common element of radical innovations.

5.3 Organisational Contingency Factors

The following section will present the results of the analysis concerning the two organisational contingency factors outlined in the introduction, namely level of organisational agility and degree of product innovativeness. Furthermore, an additional contingency factor, which focuses on the influence of industry on the relationship between BDA and NPD performance, will be presented as *ex post* results. As the choice of contingency factors only occurred after the first round of data collection, the statistics are based on a sample of seven interview respondents. The degree of product innovativeness will be explained first, followed by organisational agility. The *ex post* Industry results will be presented last to conclude the results and analysis chapter of this study.

5.3.1 Level of Organisational Agility

With regards to the organisational agility contingency factor, six out of the seven interview respondents suggested that the influence that BDA could have on NPD performance was contingent on the degree of organisational agility in the firm. This contingent relationship was positive meaning that BDA could have more influence with greater levels of organisational agility.

Shah (2019) suggested that organisational agility was crucial to BDA being able to influence NPD performance and provided an interesting metaphor as to the influence that organisational agility has on this relationship. He stated, *"if I look at data analytics as a greyhound and then you put organisational agility on top of it, that's an elephant on top of it, and tell him to run fast. That's where in my experience a lot of things die, they just don't move."* Shah (2019) noted that the influence that organisational agility has is not on the quality of the insights derived from BDA, but rather on the value of the insights for improving NPD performance, as it determines whether a company can execute based on those insights.

Building on this distinction, Dülfer (2019) dived deeper into the influence that organisational agility can have on a NPD performance criteria level. Dülfer (2019) suggested that there is *"always a positive relationship"* between organisational agility and the three performance criteria and stated, *"So with regards to time I would definitely say it's a positive relationship. With regards to quality, the same. And costs, the same."* He provided a couple of examples to demonstrate his statements. In the first example, Dülfer (2019) stated that through BDA firms gain insights which allow them to learn more about their customers and thus identify a larger variety of products that their customers may be interested in. However, in order for a company to leverage this potential and actually create these new products, which would lead to a greater product-market fit, agility is an advantage (Dülfer, 2019). The second example relates to the ability for a firm to increase its speed-to-market as a result of BDA reducing testing cycles, which was detailed in the *New Product Verification* theme above. In this example Dülfer (2019) asserted that "*now instead of spending three months on testing the product we only need two weeks and then it's done. Then we need to be also agile enough to have this next iteration, the next gate decision at an earlier point in time. I think this explains why you need this agility."* In both examples, Dülfer (2019) reinforces Shah (2019), as he states that organisational agility does not influence the quality of the insights generated through BDA but rather affects the value of the insights that are generated for improving the NPD performance. This is reasoned by the argumentation that if a company cannot execute on the insights generated then there will be no influence on NPD performance.

Through the analysis of the results it was interesting to note which are the most important underlying drivers of organisational agility for NPD. Dülfer (2019) noted that the most important area was governance, by which he meant the extent to which NPD teams are able to make decisions on their own. If this is the case then *"you are also automatically a bit more agile, you can leverage stuff in a faster way"* and therefore the *"more [NPD teams] can use big data"* (Dülfer, 2019). This point was reinforced by Antille (2019) when he stated that by developing holistic statistical tools which can be used by non-statisticians it allows them to combine their product expertise with the power of the tool and *"the two together can be quite agile to develop the products"*. This builds on Dülfer's (2019) observation as it means that now the innovation teams are able to conduct the analysis and make the decision as to how proceed themselves without consulting a data science department. Ahlbrand (2019) added that it is important for product employees to have the ability to use those analytics tools so that they do not have to continue always submitting an order to the data science team.

Whist governance was given as the most important driver, and reinforced by multiple interviews, the existence of a manufacturing setup was suggested as a reason for reducing a company's organisational agility. Thomsen (2019) stated that the influence that BDA has on NPD performance is limited when a firm has a large manufacturing setup as they are then less able to respond to identified trends. Thomsen (2019) explained that Pandora *"have to start planning 2021 now and that, of course, gives the challenge that between now and 2021, we may figure out that we thought unicorns were good but, in the meantime, unicorns have gone out of fashion."* In an ideal world, Thomsen (2019) argued, a firm would be able to recognise the changing trends via BDA and then be able to react to it but telling *“20,000 people to do something in a different way and gear such a huge operation into something different”* is extremely challenging. Therefore, Thomsen (2019) is arguing that Pandora’s large manufacturing setup reduces its organisational agility which, in turn, may mean that it cannot profit fully from its ability to use BDA for identifying new customer trends. In other words, Pandora’s low levels of organisational agility mean that even if new customer trends are identified, it would not be able to react to them, thus reducing their overall value for improving its NPD performance.

Despite Anjos (2019) agreeing in principle the level of organisational agility had a positive influence on the relationship between BDA and NPD performance, she highlighted another circumstance where organisational agility is limited. Anjos (2019) stated that in Coloplast’s particular case the positive influence of organisational agility on the relationship between BDA and NPD performance was not very strong because of the regulated industry in which Coloplast operates. Anjos (2019) said that *“the agility of Coloplast and companies in the same industry is limited by the regulation.”* Therefore, it does not matter how much organisational agility Coloplast has because the bottleneck is created by how quickly the industry can adjust. The consequence of this for Coloplast is that some of the potential of using BDA for NPD is not attainable (Anjos, 2019). Anjos (2019) alludes to this when she states that *“we are optimizing part of the process, which is the data collection, […] which is actually better, it's faster. But the outcome is still the same. So, we still need to file a patent and it will take years until we get the patent approved.”* Therefore, even if Coloplast had very high levels of organisational agility, and could adjust to new insights, the regulation in the industry mean that this could hinder the influence that BDA has on NPD performance, particularly speed to market.

5.3.2 Degree of Innovativeness

With regards to the product innovativeness contingency factor, five out of the seven interview respondents suggested that BDA could have more influence on NPD performance when the products being developed represent incremental innovations. In other words, these respondents believed that the influence of BDA on NPD performance was contingent upon the product innovation being more incremental in nature. In contrast, one respondent asserted the opposite to be true, that it would be greater with radical innovations. Finally, one respondent claimed that BDA could have an influence on NPD performance regardless of the level of product innovation.

As set out in the conceptual framework the intention was not to definitively categorise innovations as either being incremental or radical and then assess the influence of BDA on NPD performance in each case. Rather, it was to explore if relative differences in product innovativeness could affect the influence of BDA for NPD performance. Consequently, the term innovativeness was operationalised on a spectrum, from incremental to radical, and not as a series of categories as per Garcia and Calantone (2002). Therefore, the responses only reflect at which end of the innovativeness spectrum BDA could rather have an influence. Highlighting this again is important because the respondents had varied strengths of opinion as to the influence that BDA could have in the NPD process for incremental and radical innovations. On the one hand, Barfort (2019) was very clear in his view that it could be more useful with incremental innovations, *"in rare, very, very rare cases it will generate radically new ideas."* Similarly convinced was Dülfer (2019) of the opposite view, that BDA is more influential for radical innovations "*I think big data is especially powerful if you want to figure out the unknown unknowns, this is exactly what I am sure about"*. On the other hand, Anjos (2019) was much less sure about which type of innovativeness would be preferable for BDA, *"I would say giving a really, really big guesstimation here, it's incremental, the majority of the cases"*. Therefore, whilst the strength of the responses varied, by putting the degree of product innovativeness on a relative scale it still allows for an indication of the potential of BDA for NPD performance with both types of product innovations.

More respondents believed that NPD projects focussing on the incremental end of the innovativeness spectrum can profit more from using BDA than radical ones. Analysing the responses, there is a common theme running through this thinking, namely, that in order to conduct BDA successfully it is necessary to have a certain volume of data. Given this basic understanding, the interviewees suggested that as companies have more data about their existing products, then this leads to more BDA related to existing products which encourages firms to pursue more incremental innovation (Ahlbrand, 2019; Anjos, 2019; Thomsen, 2019). As Ahlbrand (2019) stated *"you need a certain amount of basic data if you're doing data analysis on old machines. If you then want to make an improvement, you can build on the old data and see what we can improve here. And that's why we're still quite clear about [pursuing] product enhancements at the moment"*. Enhancements of existing products are a typical example of incremental innovation. Further, Anjos (2019) stated that because customers *"are dealing with your product and they are usually giving feedback about things that don't work or they wish could be better"* then it is easier for companies to understand these needs and improve on them rather than innovating radically new products.

In turn, Barfort (2019) stated that the problem with completely new tangible product designs is *"that usually we have very little data"* and so it is hard to see how BDA *"will fit naturally into that process"*. He expanded on the point above and stated that as radical product innovations involve creating a new technology and a new market, there is always an element of innovating to meet currently unmet needs. However, the issue is that *"it is very hard to get data about unmet needs"* which can guide the NPD process (Barfort, 2019). Barfort (2019) illustrated this fundamental issue related to radical innovations as follows:

> *"So, designing a new banking app requires fundamentally understanding what people need today that aren't being met by current offerings. And that is, again coming from like the more data side, that is fundamentally a missing data problem. We don't have the data to be able to identify what needs are unmet. Because, you know, I can't download that data anywhere. […] So, when you're trying to basically build something that's radically new, I'm more sceptical of how data analytics can be used to generate and to gain insights into needs that aren't met at the moment."*

In this context, Baiyere (2019) concurred with the sentiment expressed above when he said that BDA *"only helps so far but I think the human intuition, the ability of the humans to imagine, you can't replace that with the analytics."* Here Baiyere (2019) was alluding to the fact that, as set out in the example of the banking app, radical innovations represent a departure from the current status quo and so this requires an element of imagination of different potential scenarios, for example, relating to the adoption of a new technology associated with a radical innovation. However, because BDA *"relies on using the past to give you insights"* it is limited in its ability to imagine scenarios which is necessary for radical innovation. Barfort (2019) reinforced Baiyere's logic by providing a converse example, where he said that by pursuing incremental improvements companies are *"able to, based on broad trends and behaviour today, get a pretty decent idea about what people are interested in using in six months or a year."* However, both respondents agreed that this is simply not possible when designing new products that represent more radical product innovations.

Summarising the statements set out above, BDA fundamentally requires a certain volume of data in order to generate insights which can be used in the NPD process and companies have more data for incremental product innovations as this, for instance, can come from their own products (Anjos, 2019). In contrast, as radical innovations tackle unmet needs there is inherently less data to use for

BDA in the NPD process (Barfort 2019). Therefore, the interview partners suggested that the low volume of data associated with radical innovations mean that the quality of the insights generated from BDA are lower, thus ultimately the influence on NPD performance is also lower. In contrast, the opposite is suggested to be true for incremental innovations.

Even though the majority of the respondents suggested that BDA can have a greater influence on NPD performance with incremental product innovations there were also some interesting arguments in favour of radical innovations. As outlined previously, Dülfer (2019) suggested that by its very nature BDA can help to create radical innovations as it is particularly suited to identify the *"unknown unknowns"*. By this, Dülfer (2019) cited the fact that BDA can help firms to *"identify what customers want that they are not aware of"* so that when compared with more traditional analytics it is possible to *"learn way more about people without them even mentioning it"*. Dülfer (2019) was implying that BDA allows companies to gain insights based on behavioural data which allows them to improve NPD performance with radical innovations. Following the same logic, Andersen raised the oft-cited anecdote that if you asked consumers 15 years ago whether they wanted an iPhone, representing a radical innovation, they would have been not sure. In contrast, behavioural data gives a new edge to NPD, as you do not have to rely on "*talking, talking, talking because quite frankly [customers] never know what they want*."

On a slightly different tangent, Baiyere (2019) stated that BDA *"would be more present in a larger proportion of incremental innovations than in radical innovations"* however this is due to a different line of reasoning. Baiyere (2019) made the interesting, and accurate, point that *"incremental innovation is what most innovation management is about"* whereas *"radical innovations are not things that you just stumble across every day"*. Therefore, purely due to the larger absolute amount of incremental innovations a firm pursues then BDA is likely to have more influence on NPD performance in this regard than with radical innovations. Whilst, the aim of this research question was to explore the influence on a case by case basis, not in absolute terms, the point raised is interesting as this may have influenced, even subconsciously, the perceptions of the other interviewees.

5.4 Ex-post Results

5.4.1 Industry

In addition to the described organisational contingency factors, which were identified on the basis of the literature review and the first five explorative interviews, there was another factor which was frequently mentioned, especially during the second data collection phase. This factor is the *industry* of the respective company and can therefore not be described as an organisational contingency factor but rather as a contextual one.

Andersen (2019) stated that every single industry will be subject to change due to the influence of BDA on the NPD process. Despite this general trend towards an increase in using BDA for NPD in all industries, many different interview partners put into perspective that the relationship between BDA and NPD performance can be influenced by the industry companies are operating in (Anjos, 2019; Vas, 2019; Antille, 2019; Thomsen, 2019, Shah, 2019; Dülfer, 2019). Interestingly, the industry characteristics that seem to be decisive for whether an industry has a positive or negative influence on the relationship between BDA and NPD differed across the conducted interviews. Generally, three different characteristics could be identified which either affect the quality of insights you can derive through BDA in the respective industry or, in contrast, the value of those insights for improving NPD performance.

With respect to the former, many interview partners highlighted that the relationship between BDA and NPD performance might be different for *digital industries* than for *non-digital ones*. In terms of the differences between non-digital and digital products, Anjos (2019) stated that regarding the use of BDA for NPD "*it's a different game*", as companies with digital products would have a clear advantage. In this context, the availability of data in particular seems to play a role in determining the quality of insights that can be derived from BDA and can thereby ultimately influence NPD performance. Companies from digital industries such as Apple, Google and Alibaba have a cleaner data set, which gives these companies a head start, indicating a higher veracity of data in digital industries (Vas, 2019). Anjos (2019) explained that these digital companies were the ones that were driving BD and thereby were the first ones that created interfaces with the end customers. Having such an interface integrated within your products is crucial for NPD, as it provides a higher velocity with "*customer information pretty much in real time*", that then allows products to be tailored to customers' preferences (Anjos, 2019). An example, suggested by Shah (2019), in this context is Fitbit.

He outlined that the company has launched many new products with new features and the "*reason they can do it is because imagine that everybody is wearing a sensor"* providing the company with high volumes of data about how customers interact with Fitbit products. With respect to such a digital interface Shah (2019) stated:

> *"So, for me that's the fundamental layer which comes first in terms of ability for your products to send you data and then you have the capability that if you have received the data then you can do all your product innovation in terms of processing it and bringing new features in and so forth."*

With regards to testing, Barfort (2019) stated that using BDA in product testing is much more challenging if products do not have some form of software component, as then customers oftentimes need to self-report their feedback, making it much harder to scale testing for non-digital products. To sum up, the interview partners suggested that by having a higher volume, a higher velocity and better veracity of data the quality of insights and ultimately the influence on NPD performance that can be achieved is higher for digital than non-digital product companies.

While this relationship in general appears to exist, it was also stressed that the disadvantage of non-digital product companies does not mean that they cannot find other ways to obtain data valuable for the NPD process and thereby derive high quality insights. Whereas digital products have interfaces within their products, non-digital product companies can build up interfaces which are external to the product. An example was given by Anjos (2019), who stated that Coloplast has launched a companion app, which supports patients using Coloplast products through their treatment. The patients can interact with the app on a daily basis and can take pictures of their stoma so that the app can identify if it is inflamed or not. Collecting and analysing such clinical data can then be valuable for developing new non-digital products, as Coloplast knows their customers and their problems better and can thereby ensure a better product-market fit (Anjos, 2019).

In addition to the degree to which an industry is digital or non-digital, another mentioned characteristic was the regulatory nature of the industry. This is outlined by Anjos (2019) who affirmed that the relationship between BDA and NPD performance is contingent on the degree of regulation in different industries. Specifically, the potential of using BDA in the Product Development stage is limited due to the Medtech industry's regulated nature (Anjos, 2019). Anjos (2019) explained that "*if* *you want to test a product in person, you need to have in Europe what we call a C-mark or in the US, what we call a medical device listing*", which restricts the testing to specific user groups. Anjos (2019) stated that this is why the scale on which you can test new products and thereby the insights that you derive from BDA about these products are always a bit limited.

While the two above mentioned characteristics influence the quality of the insights that can be derived through BDA, another characteristic was identified that determines whether the insights derived from BDA have a high or low value for improving NPD performance. In particular, it was mentioned that the degree to which customer preferences within an industry can be considered as dynamic plays a role when looking at the influence that BDA can have on NPD performance (Thomsen, 2019). Thomsen (2019) highlighted that the ability to continuously identify trends via BDA is more valuable in fast-changing environments than in more stable environments. She explained that "*being a fashion company, [Pandora] is more reliant on being able to respond to trends than for example, pharma, or Arla working with dairy and these companies*", but in turn less than fast fashion companies like Zara, which "*have to be very trendy*". According to Thomsen (2019), these fast fashion companies need to continuously identify trends to maintain a high product-market fit and hence deriving insights based on real-time data is crucial. In contrast, as Pandora is not a fast-fashion company she explains that having such continuously updated insights about the changing customer trends based on real-time data is not as valuable as for other companies because in jewellery the preferences do not change that fast (Thomsen, 2019).

6 Discussion

In the following, the results presented in chapter 5 are interpreted and discussed based on the research questions stated in the introduction to this thesis. In addition, further issues of interest that arose out of the results will also be highlighted and the implications of the collective findings will be outlined, for both theory and practice. To conclude the discussion chapter, the limitations of the study will be identified and consequently some avenues for further research, that could be of interest to pursue, will be provided.

6.1 Discussion of the Research Findings

6.1.1 Different Means by Which to Incorporate BDA into NPD

The analysis in section 5.1 identified 11 themes, which represent means by which BDA can be incorporated into NPD. There were four themes identified in Ideation and Product Development respectively, and only three in Product Launch. Whilst this appears to be an even divide between the stages, it masks the underlying weighting of the results towards the Ideation stage. Out of the 26 use cases that were provided, 13 were relevant to one of the themes within Ideation while the other 13 were split evenly across the two other stages. At first glance, such a large discrepancy between the stages would appear surprising. However, upon reflection, existing studies which cover the use of BDA for specific activities in the NPD process are overwhelmingly related to the Ideation stage, as outlined in subsection 5.1.1. With this perspective in mind it could be suggested that this thesis reinforces the trend in the literature which suggests that BDA is particularly applicable in the Ideation stage. However, here the distinction between the themes and the use cases is important, as the themes represent categories of use whilst the use cases themselves are individual examples of how BDA is used. Therefore, interpreting the findings this may suggest that BDA is equally applicable to each of the stages of the NPD process, as the themes are evenly split, whilst the weighting of the use cases reflects the actual use or popularity of BDA in each stage.

If the number of use cases in fact reflects the current popularity versus the general applicability of BDA then the paucity of use cases in the development stage could be due to the current level of development of BDA which is applicable for this stage. A key technology which was suggested as being relevant for use cases in the Product Development stage were generative modelling and generative design. Whilst these technologies do currently exist and its stated potential is accurate, Hede (2019) noted that *"it's going to take a couple of years"* before they are more widely used. Therefore, the immaturity of the technology and techniques could be a reason as to why there were less use cases identified in the Product Development stage compared to the Ideation stage. This would argue against the notion that BDA is fundamentally less applicable for the development stage.

The potential of using BDA for the Product Launch was especially highlighted by Dülfer (2019), who stated that the technologies that can be used within the activities in this stage are so advanced that they are already used on a broader scale. Dülfer (2019) and Barfort (2019) emphasized that leading media agencies show what is already possible by using BD, as all their clients' marketing mix and channel decisions about how to reach their customers in the best way is already being done via BD and being monitored on a continuous basis.

However, this lack of development in BDA technology cannot explain the relative paucity of use cases in the Product Launch stage as Dülfer (2019) and Barfort (2019) stated that the technologies for this stage are so advanced that they are already used on a broader scale, such as in marketing agencies. Therefore, the second explanation could be that the study sample is biased towards interviewees with profiles related to product development and innovation, with none of the interviewees working in a marketing function in a product company. As noted in the definition of Product Launch in subsection 3.1.2, the nature of the tasks undertaken during this stage require more marketing capabilities (R. Adams et al., 2006). Therefore, the weighting of use cases in favour of Ideation could reflect a bias in the sample.

If this were to explain the weighting then this would have wider implications for the NPD process as a whole as it would suggest that the process is somewhat siloed with innovation-related profiles working in Ideation and Product Development, and then marketing-related profiles in the Product Launch. Thomsen (2019) alluded to this issue when asked if there is some form of collaboration between marketing and innovation teams to share insights, she replied *"Not so much at the moment. It's a very new team, I would say, but I think that's a challenge that many companies have, knowledge sharing across departments"*. Antille (2019) echoed this issue as, when asked about the Product Launch phase, he did not provide any insights and justified this by saying *"we are really purely dealing with R&D and this [Product Launch] is more a marketing aspect"*. However, the literature covering integration in the NPD process is clear that cross-functional integration, especially between marketing and R&D, is a widely regarded key success factor in NPD (Ernst et al., 2010; Henard & Szymanski, 2003; Troy et al., 2008). Therefore, if this siloing exists to such an extent that the innovation profiles do not know what happens in the marketing-focussed Product Launch stage then this could reduce the knowledge sharing in the NPD process thus affecting its overall effectiveness in creating new knowledge for product development (Griffin & Hauser, 1996).

Blurring of the NPD boundaries

The development of technology which underpins BDA could also be leading to more fundamental changes in the NPD process. As stated in the literature there has been an adaptation of the traditional NPD processes, such as stage-gate, so that they include fewer stages and allow for more iteration between stages (N. Bharadwaj, 2018; Cooper, 2014). The current status quo from the interviews appears to be in line with the literature and includes three stages in the NPD process, along the lines of Ideation, Product Development and Product Launch. Anjos (2019) stated that for her the NPD process consisted of even fewer stages, only two, which she termed, Ideation and Execution. This study's findings indicate that new technology could lead to the stages in the NPD process becoming increasingly blurry. Whereas traditional NPD processes have been characterised as having discrete stages with quality controls and stop/go decisions (Jenkins et al., 2006) the stages are becoming less discrete with more overlap of activities between the stages. This finding goes beyond the current BDA literature, which has thus far merely predicted that there would be major changes in the NPD process (Johnson et al., 2017), by suggesting what these changes may look like.

Two use cases highlight this trend. Firstly, the *Fashion Preferences* use case proposed by Dülfer (2019) has been placed within the *Ideation* stage, however the output of the machine learning algorithm would be the full product design ready for testing. Therefore, this use case overlaps the Ideation stage and the early Product Development stage which includes product design. Secondly, the *AI Beer* use case proposed by Shah (2019) involves AI continuously iterating on prototypes of beer. This use case shows that the activities within the Product Development stage of the NPD process, namely product design and testing, have been merged which leads to a more fluid process that is in line with the literature. However, the underlying principle of this use case, which is that AI has the potential to autonomously design and test products without human intervention, suggests that this trend could continue. Such a proposition is in line with the study by Bharadwaj (2018) which outlines a taxonomy of strategic decision-making approaches for innovation in data-rich environments. In his study, Bharadwaj (2018) highlights the emergence of 'Algorithm-based decision making' which is where an algorithm receives salient information, evaluates the possible alternatives and recommends a required response. This method would automate much (if not all) of the decision-process and so require less managerial attention.

The consequence of this trend is that the NPD process could become increasingly integrated, as BDA-enabled activities transcend the traditional stages. This is especially interesting given the potential finding presented above that departments within the NPD process are still siloed and do not collaborate across stages. In other words, the paradox is that technologically speaking, the NPD process is becoming more integrated with increasingly blurry stages, whilst from a human capital perspective, teams are still structured according to discrete stages and thus are to some extent siloed from each other.

On the one hand, it is possible to see how this technological advancement can mitigate the effects of siloed NPD stages as the process becomes increasingly automated, thus requiring less managerial attention in the process (N. Bharadwaj, 2018). On the other hand, there is doubt as to whether decisions for innovation can be fully automated so that no managerial attention is required. Vas (2019) was also in doubt which is why he stated, *"that is why we always have what we call a human last mile, so any decision the computer takes, someone has to take it the last mile."* If this perspective is to be believed, then the technological developments may mitigate some consequences of the knowledge silos but the importance of cross-functional integration for NPD success still holds.

6.1.2 Influence of Incorporating BDA on NPD performance

The findings outlined in section 5.2 show that in general incorporating BDA has the potential to influence all three NPD performance criteria, namely cost of development, speed-to-market and product-market fit. As presented, when looking at the three criteria independently, the influence of this link can largely be considered positive.

The three criteria were selected to cover separate dimensions related to the NPD process, namely the input, process and output. However, the results showed that for many interviewees the cost of development was almost completely dependent on the time dimension of the process. Shah (2019) summarized that by saying "*many [people] when [they] think about the cost [they] always think about the time".* When being asked if incorporating BDA in the NPD process can influence cost of development, Anjos (2019) stated that she "*would believe so, because you would probably make the whole process faster*." The same point was made by Thomsen (2019), who answered to the question about what benefits an improvement in time would have for Pandora, that this would "*of course have a cost benefit*". The close link between cost of development and speed-to-market also becomes apparent when looking into the means by which BDA can influence the respective criteria. For both criteria, the findings revealed that BDA can have an effect through either shortening existing activities, substituting existing activities or eliminating whole iterations and thereby wasteful actions. Although some use cases also exemplify a reduction in cost of development by substituting traditional approaches with BDA-based approaches requiring less costly input factors (Ahlbrand, 2019), this study's findings generally question in how far BDA can actually reduce the cost of development apart from increasing the speed-to-market. This lack of clarity also calls into question the extent to which the criteria represent a suitable choice for investigating NPD performance.

Moreover, it was outlined that the positive influence of BDA on the cost of development and speed-to-market is mainly existent when looking at the relationship over a longer timeframe. The findings revealed that, especially in the short term, incorporating BDA can actually increase the costs of development through investments in set-up costs (Shah, 2019) or to build dashboards (Ahlbrand, 2019). Further, it can also decrease the speed-to-market due to the existence of certain learning processes that will naturally go along with a transformation towards more data-driven approaches (Ahlbrand, 2019). This finding is in line with the literature, which states that companies need to make considerable investments in their BDA initiatives and that these investments "*may not start yielding the desired results immediately*" due to the novelty of BD and its related technologies (Gupta & George, 2016, p. 1052). In this context it was mentioned that these initial investments, such as investments in BDA models, generally are not for one specific NPD project and so the costs can be spread across multiple NPD projects in the future (Ahlbrand, 2019; Antille, 2019). However, the benefit of spreading the costs is mitigated by the fact that as new data types advance or business conditions change these models must be adjusted which will inevitably involve extra costs (Baesens et al., 2016).

The purpose of this thesis was to explore the general potential to reduce the costs of development and speed-to-market by incorporating BDA into the NPD process and therefore quantifying these potential improvements is outside of the scope of this work. Having said that, the authors nonetheless have identified two specific issues that arise when trying to do so, which merit further attention. First, the findings of this thesis raise the question of how long should the timeframe be to evaluate cost and speed improvements? The indication in the literature is that this time horizon should be *"sufficiently long and appropriate"* (Baesens et al., 2016, p. 815), which unfortunately does not considerably improve clarity around the issue as it signals that this judgement should be made on a case-by-case basis. Second, it is questionable as to how far companies can already foresee the number of new NPD projects for which the model or dashboard can be reused. Estimating the number of projects for which models are suitable is especially difficult, as these models have to be adjusted when business conditions or BD technology changes.

These issues are particularly critical, as the purpose of these analytical models should be to create an economic return, either by raising profits or cutting costs, which should be taken into consideration by managers when making investment decisions about incorporating BDA (Baesens et al., 2016). The need to assess the economic return of incorporating BDA in different activities in the NPD process is also highlighted by Baiyere (2019) when he stated that "*in some cases you don't kill a fly with a bullet"* thereby alluding to the fact that from a cost and speed perspective it is not always worth incorporating BDA. Hence, it can be assumed that these difficulties to quantify the cost and speed improvements of incorporating BDA into the NPD process are one of the reasons why BDA is still mainly used for operations, where these improvements are more easily quantifiable.

With respect to product-market fit, the findings were mainly positive, especially for the incorporation of BDA in the Ideation stage, as was highlighted in the *Trend Spotting* and *Customer Preference Identification* themes. As product-market fit refers to creating a product that meets the requirements of the market (Schilling, 2013), it becomes clear that identifying relevant trends (Dülfer, 2019), distilling key attributes (Thomsen, 2019), or revealing suitable sub-segments for specific features (Shah, 2019) through BDA can be valuable to improve this criterion. In this context, the use of community data and subsequent social media analytics was highlighted as having great potential (Liu & Kop, 2016). This is in line with literature, which outlines that through social media analytics, insights can be derived that can lead to successful product launches (Moe & Schweidel, 2017). Although this emphasizes the potential of social media analytics, initial concerns were raised during the interviews as to its efficacy in improving the product-market fit. Barfort (2019) attributed his concerns to a lack of granularity in the data, whilst Blom (2019) highlighted the risk of biased data when he stated that *"if you feed the algorithm biased data like that then their recommendations or clusters that it will come up with […], they wouldn't be correct."* The concern around biased data being inputted into BDA is echoed in the literature. Moe and Schweidel (2017) question the representativeness of the target population in social media because not all consumers will generate content, leading to a risk of a bias within community data. This indicates that when biased data is used to make inferences about potential customers the product-market fit and thereby the failure rate of new products is potentially not improved.

While community data is freely available, which of course is advantageous for the speed and costs of data collection (Moe & Schweidel, 2017), this creates a caveat with regards to the product-market fit. Given that every company can access community data, competitors will have the same chance to identify trends and preferences that are inherent in the data. This is especially interesting given Schilling's (2013) definition of product-market fit which says that products not only need to fulfil customer requirements, but also should do so better than the products of your competitors. With this definition in mind it raises the question of how far simply having access to data allows for successful differentiation from competitors, in particular if it is publicly available data. Vas (2019) noted that accessing the data is only one relevant aspect, but that subsequently companies also need to process the data which may involve building a custom model to analyse the data. This indicates that when the first part of the BD value chain (i.e. the generation of data) is similar or even identical, then in order to derive better insights than competitors, companies need to have comparatively better models to analyse the data. Vas (2019) summarized the logic as follows:

> *"So, how do you beat the competitors? The data is the same. Everybody is looking at the customers. And this is where modelling of data, machine learning, comes into picture. So how do I beat my competitor by 5, 10 or 15%? So that [modelling of data] is the core differentiator."*

However, this statement is built on the assumption that models are proprietary, but the extent to which this will be the case in the future is questionable. Dülfer (2019) highlighted that there are already a few ready-to-use BDA tools, such as *Scatterblogs*, that can be used to, for instance, identify trends in community data. In general, this suggests that incorporating BDA can help companies to match their customers' preferences and requirements, but that in order to do so better than competitors requires them to differentiate themselves along the BD value chain.

The last paragraph highlighted that with community data the techniques employed to analyse such data are an important factor for differentiation. However, more and more companies have built interfaces to receive proprietary data from their customers. Interview respondents highlighted two examples of such interfaces, namely Coloplast with their companion app (Anjos, 2019) or Nestlé with sensors in coffee cups (Antille, 2019). Andersen (2019) described such proprietary data, which is directly related to a company's own customers, as the most important data source for innovation. Given that this data is private company data, it is likely to be better suited to allow companies to identify customer preferences that might not be apparent in their competitors' data sources. The underlying logic behind the advantage of private data over public data stems from the VRIN framework within the resource-based view (Barney, 1991). This framework states that a firm can only achieve a sustained competitive advantage if it possesses resources which are "*valuable, rare, inimitable and non-substitutable"* (Barney, 1991, p. 117). Therefore, public data would, at the very least, contravene the rare and inimitable elements as companies could easily imitate the resource and as it is free it would also not be particularly rare. Moreover, this data is derived directly from customers and can oftentimes be generated based on actual behaviour, such as in the outlined cases above. Hence, it can be assumed that there is a lower risk of bias compared to community data and so represents a more reliant source of data to improve product-market fit. Regarding behavioural data Andersen (2019) made the point that "*data does give a new edge to that, that you don't only have to rely on talking, talking, talking because quite frankly [customers] never know what theywant".*

Concluding, it can be said that the analysis of BD, especially behavioural data, can definitely improve a company's ability to identify trends and requirements and can thereby help to build products that better match customer preferences. However, the extent to which BDA can also help to match these preferences better than competitors, thus differentiating them, will depend on how companies distinguish themselves in sourcing and analysing data.

The NPD criteria Trade-off

While the discussion above considered the influence on the performance criteria independently, it was also suggested that it is important to assess the influence of BDA on the criteria collectively. Against this backdrop, Dülfer (2019) mentioned that in some areas BDA can improve one of the criteria, while at the same time worsening another. He stated that in every project, companies balance quality, cost and time which is also valid when incorporating BDA into NPD projects. While there are many potential trade-offs arising when making decisions about whether and how to incorporate BDA into the NPD process, one specific example that can be discussed in detail is whether BDA can be used as a substitute or as a complement in the NPD process. Particularly in the Ideation stage, the findings have shown that BDA is seen as a complement to traditional activities such as design thinking workshops or surveys (Andersen, 2019; Baiyere, 2019; Vas, 2019). The reason why BDA is not seen as a substitute is that it would only provide half of the picture, and only by direct customer interaction could the full picture be ascertained (Baiyere, 2019). This is in line with the literature, which states that BD can mainly explain *what* customers have done but not *why* (Hofacker et al., 2016). Therefore, although activities, such as collecting *"little data"*[5] through surveys, are more time-consuming, they are also better suited to derive attitudes, orientations or intentions (Johnson et al., 2017). This line of reasoning indicates that substituting the whole front-end of the NPD process with BDA-based approaches could potentially improve the speed-to-market and reduce the cost of development, especially when using free data sources (Chan et al., 2016). However, it might also increase the risk that product-market fit is worsened, particularly in light of the considerations above regarding the caveats of using community data.

In the light of the trade-off explained above between speed and cost on the one side and product-market fit on the other, it is interesting to consider one specific use case, mentioned in the *Customer Preference Identification* theme. The use case states that fashion companies are using machine learning to directly create the product design based on preferences identified via community data, without any direct customer interaction (Dülfer, 2019). In essence, this use case exemplifies a company that has substituted their entire front-end of the NPD process with a BDA-based approach. Whilst the literature would say that this is not the most optimal method of integrating BDA in the NPD process the reality is more nuanced and appears to depend on two factors, namely, a company's strategy and its context. First, it could be assumed that it is strategically important for a fashion company to prioritise speed-to-market, as in the fashion industry the ability to move fast and respond to upcoming trends with new products is of high importance (Thomsen, 2019). Therefore, a company could decide to prioritise speed-to-market over product-market fit as part of its strategy. Second, platforms such as Instagram represent data-rich environments for fashion companies, and higher volume, variety and quality of data from such platforms reduces the need for direct customer interaction (Vas, 2019). Therefore, the context of the company in this use case means that the deterioration in the product-market fit is not so strong and remains at an acceptable level. Concluding, the use case exemplifies that while BDA can improve each of the performance criteria individually, companies will be challenged by trade-offs in which they have to reflect on the relative importance of

[5] *Little data* is the term used to distinguish from *Big Data*

each of the criteria in their company context and make decisions to incorporate and leverage BDA in the NPD process accordingly.

Thus far, this discussion has reflected on the *potential* for each criterion and also on the inherent trade-off between the three criteria. However, this thesis has barely touched upon the organisational barriers that play a role in the context of incorporating BDA in the NPD process and which might prevent companies from realising the potential of BDA (Gupta & George, 2016). The relevance of these organisational barriers was also highlighted through the interviews. For instance, Thomsen (2019) outlined a cultural barrier, stating that not everyone trusts BDA and, especially in the case of Pandora, designers are not very interested in data, which makes it harder to realize its full potential. Similarly, the previously mentioned issue that NPD teams from distinct departments are not collaborating across stages can be seen as another organisational barrier to realise the full potential of BDA for improving NPD performance (Thomsen, 2019; Antille, 2019). Such a lack of collaboration and knowledge sharing could potentially lead to additional unnecessary analysis in the latter stages of the NPD process, which could, in turn, have a negative effect on both the cost of development and speed-to-market. Although these considerations appear important when reflecting on the potential of BDA in NPD, they are not considered part of the BDA research area, rather they are located in the Big Data Analytics Capability (BDAC) research area and thereby outside of the scope of this thesis. The concept of BDAC essentially extends the view of BD and BDA to all associated organisational resources that are important for harnessing large amounts of data to its full strategic potential (Mikalef et al., 2017) and is defined as the *"firm's ability to assemble, integrate, and deploy its big data-specific resources"* (Gupta & George, 2016, p. 1049). However, the authors considered it to be important to acknowledge the existence of such barriers and to highlight their potential influence on the degree to which the potential can finally be realised in companies.

6.1.3 Level of Organisational Agility

The results presented in subsection 5.3.1 suggest that the influence that BDA can have on NPD performance is indeed contingent on organisational agility and that this contingency is positive in nature. In other words, BDA has more influence on NPD performance the greater the levels of organisational agility in the firm. Notably, the contingency effect is not on the quality of the insights generated from BDA for NPD but rather on the value of the insights derived from BDA to improve NPD performance.

Whilst there is no literature which directly analyses the same relationships as they have been conceptualised in this thesis, it is worth contextualising the findings of this study by comparing them to the findings of other studies involving concepts similar to BDA and organisational agility. Mikalef et al. (2019) find that dynamic capabilities mediate the effect of BDAC on incremental and radical innovation capabilities and that this effect is significantly positive in both cases. Furthermore, Shuradze et al. (2018) suggest that organisational agility positively mediates the relationship between MDAC and innovation success. Although none of these studies are analysing the exact same relationships as this thesis, their findings echo those of this study which suggest that higher levels of organisational agility are beneficial for the use of BDA in NPD.

As set out in the introduction this thesis aimed to explore *how* the relationship between BDA and NPD performance is contingent upon the level of organisational agility. Whilst the findings generally echo the literature, the following discussion will highlight some of the irregularities and areas of interest that arose through the interviews which, it is hoped, will extend the understanding of organisational agility in the context of BDA for NPD.

In the results, in subsection 5.3.1, it is described that, according to Anjos (2019), industry regulation limits a company's organisational agility as it sets the rules of the game for a given industry. Therefore, the implication is that with high levels of regulation there is no incentive for a company to invest in its organisational agility as it will not receive any benefit, relative to its competitors, for doing so. Effectively, this creates a situation whereby the influence of organisational agility is superseded by the influence of the agility in the industry which, in this case, is reduced by the level of regulation. It is interesting to highlight the distinction that regulation sits at the industry level rather than the organisational level, as with the other mentioned factors of governance or extent of a manufacturing setup. Therefore, the influence of Industry on the relationship between BDA and NPD performance, which has been highlighted through *ex-post* results in section 5.4, is also visible through its impact on organisational agility.

Another interesting point, raised by Barfort (2019), is to distinguish between a company's ability to be agile (i.e. the levels of organisational agility) and a company choosing to be agile (i.e. the use of that organisational agility). With the former Barfort (2019) stated that the level of organisational agility that a company possesses *"is almost always a good thing"* for the influence of BDA on NPD performance. However, with the latter Barfort (2019) argued that whilst some firms may possess high levels of organisational agility, they would make a strategic decision not to use it due to the potential to cause unintended consequences for the firm as a whole. As an example, Barfort (2019) highlighted that by using BDA a company may gain lots of insights about upcoming trends that it could pursue, which would lead to a wider range of products each aimed at a specific trend. However, the more varied the product portfolio, the higher the potential that the different products could start conflicting with the main brand (Barfort, 2019). In particular, this could create problems for companies that have a *"well-functioning business and a brand that is really valuable"* and in these cases operating with less agility is preferred (Barfort, 2019). Whilst for other companies that do not possess such a strong brand they *"have to innovate very quickly or else they are having difficulty surviving"* (Barfort, 2019) and so for these companies both possessing and subsequently using organisational agility is crucial.

As highlighted in subsection 3.5.1, the concept of organisational agility suffers from a lack of clarity in the literature with a number of authors referring to a similar concept by various names. Furthermore, organisational agility in practice is also a concept which is not very tangible, unlike incremental innovations for example, and as such is not so easily comprehensible. The consequence of this is that organisational agility can be interpreted differently by each interview respondent despite best efforts to clearly define the term. Therefore, the combination of these two factors could be the underlying reason why the results relating organisational agility in this study were not as rich as with the other variables under consideration.

6.1.4 Degree of Innovativeness

The analysis and results presented in subsection 5.3.2 suggest that the influence that BDA has on NPD performance is contingent on the degree of product innovativeness, with more incremental innovations leading to greater influence on NPD performance. Unlike with organisational agility, the contingency effect is on the quality of the insights generated from BDA which, in turn, influence NPD performance. In other words, with incremental innovations the quality of the insights generated from BDA are higher, due to larger volume of data, which in turn lead to greater influence on NPD performance. Whilst with radical innovation the opposite is true.

The finding that BDA influences the NPD process with incremental innovations is in line with the few studies that exist in this field (Johnson et al., 2017; Mikalef et al., 2019). However, the literature also suggests that BDA should influence NPD performance positively with radical innovations (Mikalef et al., 2019), whilst the findings of this thesis raise doubts about this relationship. Given that this goes against the, albeit limited, literature on the topic it is interesting to discuss the underlying motivations behind these assertions. There were two reasons given as to why radical innovations mean that BDA is not as influential on NPD performance as incremental innovations.

One reason, put forward by Barfort (2019), was that as radical innovations involve a discontinuity along both technological and commercial axes (Garcia & Calantone, 2002), there is always an element of innovating to satisfy as yet unmet needs. However, the key issue, according to Barfort (2019), is that it is very hard to generate data about unmet needs. If this is assumed to be true, then it would preclude a considerable number of innovations from being suitable for BDA as it fundamentally requires data in order to generate insights. In turn, this would limit the overall applicability of the BDA concept for NPD. To illustrate this point, if the assumption on the categorisation of innovations was relaxed to include 'really new' innovations, which were outlined in subsection 3.5.2, then it is possible to see how BDA would also not be suitable for this category as these also involve unmet needs but only related to *either* a technology *or* a commercial perspective.

Building on the first issue above, then as radical innovations involve a departure from the status quo in the industry, they have less data available upon which to base decisions and therefore require an element of imagination during the NPD process (Barfort, 2019; Garcia & Calantone, 2002). However, this leads to the second reason, which is that BDA cannot replace the human ability to imagine new future scenarios, as it relies on using past events to create insights about the future (Baiyere, 2019). The issue raised by Baiyere (2019) merits a discussion around the fundamental applicability of BDA for incremental versus radical innovations.

The underlying method by which BDA influences NPD performance is by augmenting, or even substituting, the decision-making process of the human actors involved. This is in line with the literature which states that a fundamental benefit of BDA is as a tool to support decision-making (McAfee & Brynjolfsson, 2012). Specifically, BDA has the ability to reduce the bounded rationality of humans by increasing the computational power and analysing enormous data sets, which allows for the most rational option to be pursued based on a given objective (N. Bharadwaj, 2018). However, the concept of bounded rationality is based on the premise that all future options are *knowable* but humans are merely limited in their ability to process all of the information to choose the most rational option (Simon, 1955). So in fact, as Augier and Kreiner (2000) correctly state, what is bounded is not the rationality, but the computational ability of humans. Applying this logic to NPD, if the process, outcome and its subsequent market reaction are indeed *knowable* then, in theory, the use of BDA should lead to improved NPD performance. With the case of incremental innovations, which are improvements to existing technology in existing markets (Garcia & Calantone, 2002), it is possible to see how the process, outcome and market reaction could be *knowable* as the innovation is closer to a company's existing offerings and the analysed data is also closer to the current reality (Barfort, 2019). Therefore, from a theoretical perspective, the use of BDA has the potential to lead to improved NPD performance for incremental innovations, a proposition which is reinforced by this study's findings.

However, as radical innovations are discontinuities in technology and market structure on an industry level (Garcia & Calantone, 2002), it raises the question of whether the future is indeed *knowable* in these situations and whether reducing bounded rationality through BDA would actually lead to improved NPD performance. The countervailing view on the idea of bounded rationality was proposed by Shackle (1958), who advocates the perspective that the future is *fundamentally unknowable* and that humans make choices with regards to the future based on a series of imagined experiences (Augier & Kreiner, 2000; Shackle, 1979). It is this notion of creating future scenarios by way of imagination which seems at odds with the potential to use BDA for radical innovations. The predictions or prescriptions proposed by BDA are not (currently) created by a process of imagination. They are rather based on an underlying logic, such as an algorithm, and so it has been suggested that BDA is limited in its ability to imagine future scenarios unrelated to the past or the present (N. Bharadwaj, 2018). As Dan Ariely, the distinguished behavioural economist, noted during an interview, *"what computer systems are not good for is generating unique things and producing inferences relevant to them"* (Nadav, 2017). In other words, whereas humans suffer from bounded rationality, machines suffer from "*bounded imagination"* (Baiyere, 2019). If radical innovations are more uncertain and fundamentally less *knowable* then this would suggest that BDA, from a theoretical perspective, would have less of an impact on NPD performance with radical innovations compared to incremental innovations.

Whilst the two arguments discussed above provide compelling, and fundamental, reasons why BDA has less influence on NPD performance with radical innovations it is worth highlighting perspectives that could limit these issues. Focussing on the lack of data for radical innovations first, it is interesting to note that Dülfer (2019), Vas (2019) and Andersen (2019) all proposed that by using BDA to analyse unstructured data from IoT technologies it is possible to gain insights into the behaviour of consumers. The benefit of behavioural data is that it can reveal insights and needs that consumers are unconscious of, so the *"unknown unknowns"* about what customers like and what they do on a large scale (Dülfer, 2019). Therefore, if behavioural data is able to provide these insights then it is hereby suggested that this could be a solution to the issue as to how to generate data that help identify unmet needs for radical innovation. Now moving to the inability for BDA to be creative, it seems that this is a fundamental limitation of the technology as it is currently understood. However, Hede (2019) insightfully noted that *"we are taking a serious crack at creativity in machine learning"* and that if there was a radical innovation *within* the BDA technologies then this could *"blow a lot of doors open"* and lead to new possibilities for BDA, even with radical innovations.

Drawing the discussion threads together, there appear to be boundaries as to the use of BDA in situations where there is limited data available and so it is required to imagine and be creative. Such imagination is called upon in those situations where the future is less likely to look like the past, which is the case for radical product innovations that break the status quo. Whilst this would suggest inherent limitations of BDA when innovating around more radical product ideas, the suggestion that incorporating behavioural data could reveal insights into unmet needs provides some hope that BDA can also help NPD performance with radical innovations. Otherwise, it appears to be reliant on a step-change improvement in the BDA technology to make it more applicable for NPD.

Aside from the discussion above, it should be highlighted that the majority of the responses from interviewees concerning this contingency factor were related to the Ideation stage. Very few comments were made, either way, regarding how the degree of innovativeness of the product would impact the use of BDA in the Product Development and Product Launch stages. Such a finding in itself is interesting as it could suggest that the distinction between radical and incremental innovations is greatest during the Ideation stage and that the process becomes more similar later on. If this were to be the case then this would represent a departure from the literature on the impact of innovativeness of product innovations on the NPD process, as outlined in subsection 3.5.2.

6.1.5 Industry

The results of this study show that *industry* can be seen as a relevant contextual contingency factor in the relationship between BDA and NPD performance. In addition, three industry characteristics were identified through the interviews that determine whether an industry has a more positive or negative influence on the above-mentioned relationship. These three characteristics are the degree to which the industry can be seen as digital, the degree of regulation in the industry and lastly the degree to which customer preferences can be assessed as dynamic in the respective industry. An especially interesting characteristic to discuss appears to be the dynamism of the customer preferences, as support for this characteristic could also be identified in extant literature. The findings of this thesis have particularly suggested that BDA has a greater influence on NPD performance when the dynamism of customer preferences is high. Put simply, when trends and preferences are dynamic, BDA has a greater influence on NPD performance. Such dynamism in customer preferences was already addressed in literature and is what Johnson et al. (Johnson et al., 2017, p. 640) term *"customer turbulence"*. Similarly to this study, Johnson et al. (2017) investigated the extent to which the use of BD can influence new product revenue, and found out that this relationship is stronger with high customer turbulence. Interestingly, new product revenue can be considered an example of an NPD output criterion (R. Adams et al., 2006; Cordero, 1990), similarly to product-market fit. Thereby the evidence is strengthened that higher dynamism in customer preferences will increase the influence BDA can have on NPD performance, specifically on product-market fit.

The findings in subsection 5.4.1 indicate that the influence of BDA on NPD performance in general will be dependent on whether these industry characteristics are present in a company's industry. However, this does not mean that BDA is a niche phenomenon and only applicable in some industries. Kiron et al. (2012) found that *analytical innovators*, which are companies that can be described as the most advanced in innovating with BD, came from a multitude of different industry sectors. This mirrors the findings of our thesis, in which four product companies from diverse industries all signalled concrete potential for using BDA in NPD.

Furthermore, through the interviews a potential shift in one of these industry characteristics could be identified, namely the proportion of products which are digital. Shah (2019) pointed out that industries are going to change a lot in any case, and industries that produce non-digital products today are becoming increasingly connected. He states that *"there is one thing which is definitely happening, that products are getting connected"* and thereby all products are increasingly creating self-quantification data. Once companies include IoT technology in their products they will realize "*in which country [their products] are, what kind of operations they're doing, what kind of failures they are having. Then it becomes so valuable*". This statement indicates that the boundaries between digital and non-digital industries will continue to blur, and thus, the influence BDA can have on NPD performance is expected to rise in many industries. This trend is also validated by literature, as more and more products will became connected and by 2030 onwards IoT data will account for the largest share of BD (M. Chen et al., 2014). Even though some products themselves might not be connected in the future, for example consumable products such as milk, the existence of cameras and sensors in supermarkets and the rise of smart-home devices would allow these product companies to receive large amounts of data that are valuable for NPD (Barfort, 2019). This indicates that the richness of the BDA will increase in the future through a rise in relevant data across industries, thereby enabling companies that produce non-digital products today to derive a higher quality of insights for NPD and thereby a potentially higher influence on their NPD performance. Such a development would further strengthen the relevance of this thesis.

6.2 Implications

6.2.1 Practical Implications

This thesis sets out a number of practical implications that can be valuable for practitioners when considering whether, and how, to incorporate BDA into their NPD process. In this context, the authors want to highlight that the implications of this thesis are mainly targeted at practitioners from the field of innovation management and probably less so for practitioners from the field of data science. This is due to the fact that this thesis does not deal, in great detail, with the different BDA technologies and techniques used in the different activities during the NPD process, but is more focussed on the business potential of incorporating BDA in NPD.

First, this study provides practitioners with a detailed typology of the different means by which BDA can be incorporated into the NPD process, thus going beyond the loose expressions of BDA's potential for NPD which characterises most of the current literature in this field. In so doing, this thesis addresses the problem raised by Mikalef et al. (2018, p. 548) that practitioners *"are left in unchartered territories"* when incorporating BDA. In contrast to the extant patchwork of studies, this thesis provides practitioners with a state-of-the-art overview, offering a more holistic picture of the intersection of BDA and NPD. By involving consultants and representatives from product companies, and by developing the typology based on 26 real-world use cases, this thesis is based to a large extent on practical insights. This increases the relevance of the results for practitioners, shining a light on how other companies have already incorporated BDA in their NPD processes. Furthermore, by highlighting certain contingency factors, this thesis also underlines that the applicability of the themes and use cases are not universally valid but are contingent on at least three factors that were identified throughout this research. The findings around these three factors can give practitioners guidance into what should be considered with regards to their own company and industry, when incorporating BDA into their company's NPD process.

Second, this study highlights that practitioners should regard BDA as a complement rather than a substitute to traditional NPD activities, especially in the Ideation stage. Hence, this cautions practitioners against relying solely on the insights derived from BDA about identified trends and preferences, as they may not show the full picture and sometimes bear the risk of being biased. In contrast, incorporating BDA as a complement can provide insights that would otherwise not be possible to derive, especially when leveraging behavioural data, thereby giving NPD a new edge.

Third, the general potential of incorporating BDA has been highlighted across all three phases of the NPD process. However, companies still seem to suffer from silos in the workplace, which is reducing cross-functional integration and potentially harming the NPD process. To leverage the full potential of BDA, practitioners need to break down these structures and exploit the synergies created through more integrated BDA-based activities. As an example, the marketing department would then be able to leverage insights for the Product Launch stage, which were initially derived in an earlier stage by the product development team. This is important as otherwise unnecessary costs for additional analysis are occurring and the speed-to-market is needlessly reduced.

Finally, this thesis indicates that with the increasing adoption of BDA for NPD across industries, the use of publicly available data and ready-to-use BDA tools will not enable companies to differentiate themselves from competition. The implication for practitioners is that they need to consider ways to gather proprietary data, for instance through the creation of digital interfaces, such as apps, and develop proprietary analytical models. Thereby companies will be able to derive insights that the competition does not possess, potentially leading to a sustained competitive advantage. If these companies do not move in this direction and make the necessary investments, they will be outmanoeuvred by their competitors (Vas, 2019).

6.2.2 Theoretical Implications

With respect to the theoretical implications, this study provides further evidence of the potential of incorporating BDA in the NPD process and thereby validates the importance for further research in this emerging research area. The main theoretical contribution can be seen in the development of a typology, consisting of 11 themes that represent the overall means by which BDA can be integrated into the NPD process. The typology provides structure to this emerging research area and allows existing research to be classified according to one of the themes. Moreover, and with respect to future research, the 11 themes represent concrete fields which can guide further investigation by other scholars.

Furthermore, this study contributes to theory by emphasizing the importance of contingency factors when researching the potential of BDA, specifically when looking at NPD. Besides organisational contingency factors, contextual contingency factors were also found to be important to consider. With respect to the role that the degree of innovativeness plays for the potential of using BDA, the extant literature suggested that it is applicable for creating both incremental and radical innovations. This thesis extends the literature by indicating that, while the potential is apparent for both, the influence of BDA for NPD performance seems to be stronger with incremental innovations. Further, with respect to the role of organisational agility, the findings are in line with the literature that organisational agility is an important capability to execute on the BDA insights to generate business value. Interestingly, this thesis presents evidence that the relative importance of organisational agility is linked to industry-specific characteristics such as the level of regulation. Moreover, this study contributed to literature by stating that the degree to which BDA can influence NPD performance is highly contingent on the industry and the underlying drivers. In particular, the industry factor was found to have an influence on the ability to perform BDA and also on the value of the insights derived.

6.3 Limitations and Further Research

The following section presents the general limitations of this study. When possible, these limitations will be linked to the methodological limitations previously presented in section 2.5. Additionally, further research on how these limitations could be mitigated through future studies is outlined. Thereafter, additional avenues of further research are presented, that have been unveiled in the context of this thesis and are worth further investigation.

As outlined in the methodological limitations, qualitative research always suffers from limitations related to the sample under investigation and this study is no exception. While some limitations caused by the sample were recognised at the outset of this study and partly mitigated through various means, as described in section 2.5, other limitations only became apparent in the course of the study.

Evaluating adequacy of samples in qualitative research can refer to considerations of sample size and sample composition (Vasileiou et al., 2018). Evident from the outset was that the small sample of interview partners would limit the exhaustiveness of the typology. As the 11 themes were developed based on empirical use cases, additional interviews could have presented further use cases, which in turn could have led to the development of new themes or to the further validation of the developed ones. Hence, further research could build on this thesis by increasing the sample size in order to improve the validity and exhaustiveness of the themes.

Limitations that became apparent during the study not only refer to sample size but also to sample composition. It must be noted that the weighting of respondents towards those with product development profiles, as opposed to marketing profiles, likely impacted the prevalence of use cases in each stage. Whilst efforts were made to supplement the sample with appropriate sources, this was unfortunately not possible due to a lack of availability. Therefore, including marketing profiles in similar future studies is recommended to increase the validity and generalizability of the results.

Another limitation that emerged during the course of the study is linked to the contextual contingency factor *industry* and therefore also considers the sample composition. This contingency factor was presented as *ex-post* results because its importance was only recognised during the second data collection phase and so it was originally not part of the conceptual framework. The results of that subsection have shown that there are three characteristics that determine the influence of the industry on the relationship between BDA and NPD performance. Whilst the four interviewed product companies cover industries with different degrees of regulation and level of dynamism, none produce a digital product, which could have impacted the results by providing distinct insights, thus enabling the authors to present even more comprehensive results. This limitation is thought to have been partially mitigated by incorporating the views of consultants who have experience across a number of industries, including those with digital products.

Besides these limitations related to the sample, another limitation is linked to the theoretical concepts of BD and BDA. The authors noted that the research field around BD is increasing in popularity, with many papers providing definitions of the terms included in this study, such as BD and BDA. Despite this popularity is it noteworthy that there do not appear to be any sharp boundaries which delineate the concept of BD from just normal data or likewise BDA from data analytics. Therefore, as the authors had no clear and specific guidelines as to what would fall within the BDA concept there is the risk that the use cases could be interpreted differently thus leading to lower reliability. By selecting interviewees who were exposed to and confident with the BDA field it is hoped that this risk was mitigated. However, the authors advocate for further research that will help to clearly delineate the concepts of BD and BDA in order to provide clarity as to what is considered BD and what is not.

Moving on to the contingency factor of the degree of innovativeness, it became clear that by defining this concept on a spectrum from radical to incremental it limited the insights and did not allow the nuances to be highlighted. The assumption made regarding the linear relationship between incremental and radical innovation could very well be creating an overly simplistic distinction. Furthermore, this study assumed that all responses were equally valid, however, as noted above, some interviewees were much more confident in their responses than others when it came to the issue of product innovativeness. Therefore, this assumption could lead to biased results as a weak confirmation towards incremental innovation has been counted the same as a strong assertion for radical innovation. To mitigate this bias, it could be useful to rank the strength of responses on a scale and then assign relative weights in order to increase the validity of the results. The effect of this potential bias can be seen when viewed from the perspective that incremental innovations are more ubiquitous than radical innovations in product companies, which, in turn, could have created a subconscious bias in the perceptions of the respondents as to the applicability of BDA.

Finally, the organisational agility contingency factor was limited by the complexity of the concept which seemed to create a lack of understanding about what exactly was being asked. This was reflected in the interview responses that were either vague or even inaccurate and as such could not be included in the results for fear of reducing the validity of the research.

In addition to the further research mentioned above, which was proposed as a possible method to overcome some of the limitations of this thesis, two additional avenues of further research were identified in the context of this study that the authors consider to be worth further investigation.

First, further research could be conducted to validate the findings of this thesis with respect to the influence that incorporating BDA could have on NPD performance. The findings of this thesis indicated that BDA could positively influence all three performance criteria individually. However, due to the explorative nature of this study, these findings cannot be seen as a proven causal relationship, but instead represent indications upon which hypotheses and future research could be based. Therefore, further research should now continue along the path paved by this thesis and attempt to validate these indications through quantitative studies. Quantitative studies could hopefully provide insights into the degree of improvement in each criterion that can be achieved by incorporating BDA. Such a tangible and measurable proof of the potential that BDA can have for NPD could help to overcome the fact that it is currently still difficult to create investment cases for these types of BDA. Thus, additional studies could further encourage practitioners that investments in NPD-related BDA can yield a positive return on investment.

Second, further research should also be conducted on the effect that BD or BDA has on the NPD process. As outlined in section 3.4, literature highlighted that BDA will lead to major changes in the NPD process, or even transform the process entirely. In line with these statements, the findings of this thesis have revealed that the integration of BDA, especially in the form of machine-learning applications, leads to the blurring of individual activities and stages within the NPD process. These findings challenge the existing theory on NPD such as the stage-gate model, which differentiates sharply between different stages. Further research could clarify whether the blurring of phases is only the case for individual use cases, or whether this is a general development that becomes apparent with the increasing use of BDA within the NPD process.

7 Conclusion

While NPD increasingly faces the challenge to innovate faster and at lower cost, whilst ensuring a better fit to the market's needs, the abundance of data and the analysis of such enormous quantities of data represent a potential pathway to overcome these challenges. However, the authors of this thesis identified that the intersection between BDA and NPD, and in particular how the potential of BDA can be leveraged for NPD, to be an understudied area, hence providing little guidance for practitioners and researchers alike. Therefore, the aim of this study was to address the important questions of, first, how can BDA influence NPD performance and, second, how is this relationship contingent on organisational factors.

Due to the lack of previous research, this thesis used an inductive approach to theory development to address these questions and broadly explore the potential of BDA for NPD. In particular, the qualitative research method of Grounded Theory was chosen to iteratively refine the research focus and thereby be able to react to emerging patterns in the qualitative data that are worth further investigation. In total 12 expert interviews were held covering respondents from leading product companies, consultancies and academic institutions. The interviews were conducted in two data collection phases spanning from May until August 2019.

The findings of this thesis provide interesting insights with respect to the considered research questions. First, it addressed the question of *how can BDA influence NPD performance*. In order to answer this question, this thesis began by identifying and exploring the means by which BDA can be incorporated into NPD, through the development of a typology. Based on the conducted interviews, 26 use cases were identified, which, through a rigorous three-step coding process, were grouped into 11 overarching themes. Out of these 11 themes, four themes were identified in both the Ideation stage and in the Product Development stage whilst only three themes are in the Product Launch stage. After identifying the means by which NPD can be incorporated, this thesis sought to study how incorporating BDA into NPD can influence NPD performance, which was represented by three criteria. In this study, cost of development was taken to cover the inputs to the NPD process, speed-to-market was chosen to cover the process itself and the criteria of product-market fit was selected to evaluate the output of the NPD process. This thesis found that, in general, all three criteria can be positively influenced by the use of BDA in the NPD process. However, it was identified that the inherent trade-off between the three criteria, which affects NPD projects more broadly, must also be taken into consideration when incorporating BDA into the NPD process. Hence, companies must decide about the relative importance of each of the criteria according to their context and strategy and make decisions to incorporate and leverage BDA in the NPD process accordingly.

Second, this thesis also considered *how the relationship between BDA and NPD performance is contingent on two organisational factors*, namely level of organisational agility and degree of product innovativeness. Regarding the former, the results of this thesis have shown that BDA has more influence on NPD performance the greater the levels of organisational agility in the firm. With respect to the latter, the results suggest that the influence that BDA has on NPD performance is greater with incremental innovations compared to radical innovations. Besides these two organisational contingency factors, the industry of the respective company was identified as a contextual contingency factor throughout the course of the analysis. The evidence has shown that three underlying drivers, namely the different degrees of regulation, the dynamism of customer preferences within an industry and the degree to which the industry can be described as digital, affect to what extent BDA can influence performance improvements.

8 Sources

Adams, G., & Schvaneveldt, J. (1991). *Understanding research methods* (2nd ed.). New York: Longman.

Adams, R., Bessant, J., & Phelps, R. (2006). Innovation management measurement: A review. *International Journal of Management Reviews*, *8*(1), 21–47. https://doi.org/10.1111/j.1468-2370.2006.00119.x

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research*, *25*(3), 443–448. https://doi.org/10.1287/isre.2014.0546

Ahlbrand, P. (2019). Personal Communication - 16 July 2019.

Andersen, A. (2019). Personal Communication - 29 May 2019. Copenhagen.

Anjos, I. (2019). Personal Communication - 15 July 2019. Copenhagen.

Antille, N. (2018). Pre-read material for interview. *Food Quality and Preference*. Elsevier. https://doi.org/10.1016/j.foodqual.2017.08.008

Antille, N. (2019a). Personal Communication - 13 August 2019.

Antille, N. (2019b). Pre-read material for interview - Design of experiment with sensory data: A pragmatic data analysis approach. *Journal of Sensory Studies*, *34*(2). https://doi.org/10.1111/joss.12489

Augier, M., & Kreiner, K. (2000). Rationality, imagination and intelligence: some boundaries in human decision-making. *Industrial and Corporate Change*, *9*(4), 659–681. https://doi.org/10.1093/icc/9.4.659

Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). TRANSFORMATIONAL ISSUES OF BIG DATA AND ANALYTICS IN NETWORKED BUSINESS. *MIS Quarterly*, *40*(4), 807–818.

Baiyere, A. (2019). Personal Communication - 14 May 2019. Copenhagen.

Banerjee, A., Bandyopadhyay, T., & Acharya, P. (2013). Data Analytics: Hyped Up Aspirations or True Potential? *Vikalpa*, *38*(4), 1–12. https://doi.org/10.1177/0256090920130401

Baregheh, A., Rowley, J., & Sambrook, S. (2009). Towards a multidisciplinary definition of innovation. *Management Decision*, *47*(8), 1323–1339. https://doi.org/10.1108/00251740910984578

Barfort, S. (2019). Personal Communication - 26 July 2019. Copenhagen.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*.

Barton, D., & Court, D. (2012). Making Advanced Analytics Work For You. *Harvard Business Review*, (October), 78–83.

Bessant, J., Lamming, R., Noke, H., & Phillips, W. (2005). Managing innovation beyond the steady state. *Technovation*, *25*(12), 1366–1376. https://doi.org/10.1016/j.technovation.2005.04.007

Bharadwaj, A. S. (2000). A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation. *MIS Quarterly*, *24*(1), 169–196. https://doi.org/10.1016/j.neuron.2009.01.012

Bharadwaj, N. (2018). Strategic Decision Making in an Information-Rich Environment: A Synthesis and an Organizing Framework for Innovation Research. *Review of Marketing Research*, *15*, 3–30.

Bharadwaj, N., Nevin, J. R., & Wallman, J. P. (2012). Explicating hearing the voice of the customer as a manifestation of customer focus and assessing its consequences. *Journal of Product Innovation Management*, *29*(6), 1012–1030. https://doi.org/10.1111/j.1540-5885.2012.00954.x

Bharadwaj, N., & Noble, C. (2017). Finding Innovation in Data-Rich Environments. *Journal of Product Innovation Management*, *34*(5), 560–564.

Biehn, N. (2013). The missing V's in Big Data: Viability and value.

Blom, M. (2019). Personal Communication - 21 May 2019. Copenhagen.

Bryman, A., & Bell, E. (2011). *Business Research Methods* (Third). Oxford University Press.

Calantone, R. J., Harmancioglu, N., & Droge, C. (2010). Inconclusive innovation "returns" A meta-analysis of research on innovation in new product development. *Journal of Product Innovation*

Management, *27*(7), 1065–1081. https://doi.org/10.1111/j.1540-5885.2010.00771.x
Cambridge Dictionary. (2019). AGILITY | meaning in the Cambridge English Dictionary. Retrieved August 28, 2019, from https://dictionary.cambridge.org/dictionary/english/agility
Canuto, O. (2018). How globalization is changing innovation. Retrieved August 8, 2019, from https://www.weforum.org/agenda/2018/08/globalisation-has-the-potential-to-nurture-innovation-heres-how
Castellion, G., & Markham, S. K. (2013). Perspective: New product failure rates: Influence of Argumentum ad populum and self-interest. *Journal of Product Innovation Management*, *30*(5), 976–979. https://doi.org/10.1111/j.1540-5885.2012.01009.x
Chan, H. K., Wang, X., & Zhang, M. (2016). A Mixed-Method Approach to Extracting the Value of Social Media Data, *25*(3), 568–583. https://doi.org/10.1111/poms.12390
Chang, W., & Taylor, S. A. (2016). The Effectiveness of Customer Participation in New Product Development: A Meta-Analysis. *Journal of Marketing*, *80*(1), 47–64. https://doi.org/10.1509/jm.14.0057
Charmaz, K. (2006). Constructing Grounded Theory: A Practical Guide Through Qualitayive Analysis. In *Introducing Qualitative Methods*. SAGE Publications. https://doi.org/10.1186/s12868-016-0320-5
Chen, J., Reilly, R. R., & Lynn, G. S. (2012). New product development speed: Too much of a good thing? *Journal of Product Innovation Management*, *29*(2), 288–303. https://doi.org/10.1111/j.1540-5885.2011.00896.x
Chen, M., Mao, S., & Liu, Y. (2014). Big data: A Survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0
Chesbrough, H. (2019). Open Innovation and Avoiding the Commodity Trap. Retrieved August 8, 2019, from https://executive.berkeley.edu/thought-leadership/blog/open-innovation-and-avoiding-commodity-trap
Chesbrough, H., Vanhaverbeke, W., & West, J. (2006). *Open Innovation: Researching a New Paradigm*. Oxford University Press. https://doi.org/10.1111/j.1467-8691.2008.00502.x
Christensen, K., Nørskov, S., Frederiksen, L., & Scholderer, J. (2017). In Search of New Product Ideas: Identifying Ideas in Online Communities by Machine Learning and Text Mining. *Creativity and Innovation Management*, *26*(1), 17–30. https://doi.org/10.1111/caim.12202
Citrin, A. V., Lee, R. P., & McCullough, J. (2007). Information use and new product outcomes: The contingent role of strategy type. *Journal of Product Innovation Management*, *24*(3), 259–273. https://doi.org/10.1111/j.1540-5885.2007.00249.x
Collis, J., & Hussey, R. (2003). *Business Research Business Research A Practical Guide for Undergraduate and Postgraduate Students*. *Nature*.
Collis, J., & Roger Hussey. (2003). *Business Research Business Research A Practical Guide for Undergraduate and Postgraduate Students*. *Nature*.
Cooper, R. G. (1984). The Strategy-Performance Link in Product Innovation. *R&D Management*, *14*(4), 247–259. https://doi.org/10.1111/j.1467-9310.1984.tb00521.x
Cooper, R. G. (1990). Stage-Gate Systems: A New Tool for Managing New Products. *Business Horizons*, (May-June), 44–54.
Cooper, R. G. (1993). *Winning at new products: Accelerating the process from idea to launch* (2nd Editio). Cambridge, Massachusetts: Perseus Books.
Cooper, R. G. (2008). Perspective: The Stage-Gate Idea-to-Launch Process - Update, What's New, and NexGen Systems. *Journal of Product Innovation Management*, *25*, 213–232.
Cooper, R. G. (2014). What's Next? After Stage-Gate? *Research Technology Management*, *157*(1), 20–31.
Cooper, R. G., & Kleinschmidt, E. J. (1987). Success factors in product innovation. *Industrial Marketing Management*, *16*(3), 215–223. https://doi.org/10.1016/0019-8501(87)90029-0
Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative

criteria. *Qualitative Sociology*. https://doi.org/10.1007/BF00988593
Cordero, R. (1990). The measurement of innovation performance in the firm: An overview. *Research Policy*, *19*(2), 185–192. https://doi.org/10.1016/0048-7333(90)90048-B
Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitativCreswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. Educational Research (Vol. 4). https://doi.org/10.1017/CBO97811*. *Educational Research*. https://doi.org/10.1017/CBO9781107415324.004
D'Aveni, R. A., Dagnino, G. B., & Smith, K. G. (2010). The age of temporary advantage. *Strategic Management Journal*, *31*(13), 1371–1385. https://doi.org/10.1002/smj.897
Davenport, T. H. (2013). Analytics 3.0. *Harvard Business Review*, (5), 64–72.
Davenport, T. H., Barth, P., & Bean, R. (2012). How 'Big Data' Is Different. *MIT Sloan Management Review*, *54*(1), 1–7.
Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning*. Harvard Business School Press.
Delen, D., & Demirkan, H. (2013a). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359–363. https://doi.org/10.1016/j.dss.2012.05.044
Delen, D., & Demirkan, H. (2013b). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359–363. https://doi.org/10.1016/j.dss.2012.05.044
Desjardins, J. (WEF). (2019). How much data is generated each day? Retrieved from https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/
DTI. (2003). Innovation Report - Competing in the global economy: the innovation challenge. *Department of Trade and Industry.*, (December), 144. https://doi.org/10.1049/ic:20040517
Dubey, R., Gunasekaran, A., & Childe, S. J. (2018). Big data analytics capability in supply chain agility. *Management Decision*, *N/A*(N/A). https://doi.org/10.1108/md-01-2018-0119
Dülfer, N. (2019). Personal Communication - 16 August 2019.
Ernst, H. (2002). New Product Development (NPD) Success Factors : A Review of the Literature. *International Journal of Management Reviews*, *4*(1), 1–40. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.483.5376&rep=rep1&type=pdf
Ernst, H., Hoyer, W. D., & Rübsaamen, C. (2010). Sales, Markerting and R&D Cooperation across New Product Development Stages: Implications for Success. *Journal of Marketing*, *74*(September), 80–92.
Evanschitzky, H., Eisend, M., Calantone, R. J., & Jiang, Y. (2012). Success factors of product innovation: An updated meta-analysis. *Journal of Product Innovation Management*, *29*(1994), 21–37. https://doi.org/10.1111/j.1540-5885.2012.00964.x
Fagerberg, J., Mowery, D. C., & Nelson, R. R. (2004). *The Oxford Handbook of Innovation*. Oxford University Press.
Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*. https://doi.org/10.1145/2602574
FirstInsight. (2019). Optimize New Product Creation with Predictive Analytics.
Foss, N. J., & Saebi, T. (2017). Fifteen Years of Research on Business Model Innovation. *Journal of Management*, *43*(1), 200–227. https://doi.org/10.1177/0149206316675927
Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031
Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007
Garcia, R., & Calantone, R. (2002). A critical look at technological innovation typology and innovativeness terminology: a literature review. *Journal of Product Innovation Management*, *19*,

110–132.
Gaskell, A. (2018). How Open Innovation Can Reduce The Costs Of Innovation.
George, G., Haas, M. R., & Pentland, A. (2014). From the editors: Big data and management. *Academy of Management Journal*, *57*(2), 321–326. https://doi.org/10.5465/amj.2014.4002
Giannakis, M., & Louis, M. (2016). A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management*, *29*(5), 706–727. https://doi.org/10.1108/JEIM-06-2015-0050
Glaser, B., & Strauss, A. (2006). *The Discovery of Grounded Theory: Strategies for Qualitative Research*.
Gobble, M. M. (2013). Big Data: The Next Big Thing in Innovation. *Research-Technology Management*, *56*(1), 64–67. https://doi.org/10.5437/08956308X5601005
Griffin, A., & Hauser, J. R. (1993). The voice of the customer. *Marketing Science*, *12*(1–27).
Griffin, A., & Hauser, J. R. (1996). Integrating R & D and marketing: A review and analysis of the literature. *Journal of Product Innovation Management*. https://doi.org/10.1016/0737-6782(96)00025-2
Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004
Hagel, J. (2015). Bringing Analytics to Life. *Journal of Accountancy*.
Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations and Production Management*, *36*(10), 1382–1406. https://doi.org/10.1108/IJOPM-02-2014-0098
Hauser, J. R., & Zettelmeyer, F. (1997). Metrics to evaluate R,D&E. *Research Technology Management*, *40*(4), 32–38. https://doi.org/10.1080/08956308.1997.11671140
Hede, A. (2019). Personal Communication - 13 May 2019. Copenhagen.
Heidemann, J., Klier, M., & Probst, F. (2012). Online social networks: A survey of a global phenomenon. *Computer Networks*. https://doi.org/10.1016/j.comnet.2012.08.009
Henard, D. H., & Szymanski, D. M. (2003). Why Some New Products are More Successful than Others. *Journal of Marketing Research*, *38*(3), 362–375. https://doi.org/10.1509/jmkr.38.3.362.18861
Henderson, R. M., & Clark, K. B. (1990). Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly*, *35*(1), 9–30. https://doi.org/10.2307/2393549
Heron, J. (1996). Co-operative Inquriy: Research into the Human Condition. In *1996*.
Hofacker, C. F., Malthouse, E. C., & Sultan, F. (2016). Big Data and consumer behavior: imminent opportunities. *Journal of Consumer Marketing*, *33*(2), 89–97. https://doi.org/10.1108/JCM-04-2015-1399
Hoornaert, S., Ballings, M., Malthouse, E. C., & Van den Poel, D. (2017). Identifying New Product Ideas: Waiting for the Wisdom of the Crowd or Screening Ideas in Real Time. *Journal of Product Innovation Management*, *34*(5), 580–597. https://doi.org/10.1111/jpim.12396
Jenkins, S., Forbes, S., Durrani, T. S., & Banerjee, S. K. (2006). Managing the product development process. Part I: an assessment. *International Journal of Technology Management*, *13*(4). https://doi.org/10.1504/ijtm.1997.001670
Jezerc, G. (2018). First Insight Launches First Customer-Centric Merchandising Platform. Retrieved August 22, 2019, from https://www.firstinsight.com/press-releases/first-insight-launches-first-customer-centric-merchandising-platform
Jin, J., Liu, Y., Ji, P., & Liu, H. (2016). Understanding big consumer opinion data for market-driven product design. *International Journal of Production Research*. https://doi.org/10.1080/00207543.2016.1154208
Johnson, J. S., Friend, S. B., & Lee, H. S. (2017). Big Data Facilitation, Utilization, and Monetization: Exploring the 3Vs in a New Product Development Process. *Journal of Product Innovation Management*, *34*(5), 640–658. https://doi.org/10.1111/jpim.12397

Kiron, D., Prentice, P. K., & Ferguson, R. B. (2012). Innovating With Analytics. *MIT Sloan Management Review*, *54*(1), 1–6. Retrieved from http://sloanreview.mit.edu/article/innovating-with-analytics/

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*. https://doi.org/10.14778/2367502.2367572

Laney, D. (2001). 3D Data management: Controlling Data Volume, Velocity, and Variety. *Application Delivery Strategies*. https://doi.org/10.1016/j.infsof.2008.09.005

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data, Analytics and the Path From Insights to Value. *Sloan Management Review*, *52*(2), 21–31. https://doi.org/10.0000/PMID57750728

Lee, T., & Bradlow, E. T. (2011). Automated Marketing Research Using Online Customer Reviews. *Journal of Marketing Research*, *XLVIII*(October), 881–894. https://doi.org/10.2139/ssrn.1726055

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big Data in product lifecycle management. *International Journal of Advanced Manufacturing Technology*, *81*(1–4), 667–684. https://doi.org/10.1007/s00170-015-7151-x

Lieberman, M. B., & Montgomery, D. B. (1988). First-Mover Advantages. *Strategic Management Journal*, *9*(Summer), 41–58. Retrieved from http://www.jstor.org/stable/2486211

Lincoln, Y. ., & Guba, E. . (1994). Competing Paradigms in Qualitative Research Naturalistic Inquiry. *Handbook of Qualitative Research*, 105–117.

Liu, R., & Kop, A. E. (2016). Does Social Media Really Help? From Customer Involvement to New Product Success. *International Journal of Online Marketing*, *6*(3), 15–33. https://doi.org/10.4018/IJOM.2016070102

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition and productivity*. *McKinsey Global Institute*.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87. https://doi.org/10.1287/orsc.2.1.71

Marshall, C., & Rossman, G. (1999). Designing qualitative research. *Designing Qualitative Research (3rd Edition)*. https://doi.org/10.2307/2072869

Mason, H. (2018). How to Decide Which Data Science Projects to Pursue. *Harvard Business Review Digital Articles*, *October*(17), 1–5.

Massa, L., & Tucci, C. (2014). Business Model Innovation. In *The Oxford Handbook of Innovation Management* (pp. 420–441).

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68.

McCardle, M., White, J. C., & Calantone, R. (2018). Market Foresight and New Product Outcomes. *Review of Marketing Research*, *15*, 169–203.

McCracken, G. (1988). *The Long Interview: Qualitative Research Methods*. *Sage Publications*. SAGE Publications. https://doi.org/h61.28.m37 v.13

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2018). COMPLEMENTARITIES BETWEEN INFORMATION GOVERNANCE AND BIG DATA ANALYTICS CAPABILITIES ON INNOVATION.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management*, *30*(2), 272–298. https://doi.org/10.1111/1467-8551.12343

Mikalef, P., Framnes, V. A., Danielsen, F., Krogstie, J., & Olsen, D. (2017). Big Data Analytics Capability: Antecedents and Business Value. In *PACIS 2017 Proceedings* (p. 136). Retrieved from http://aisel.aisnet.org/pacis2017http://aisel.aisnet.org/pacis2017/136

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a

systematic literature review and research agenda. *Information Systems and E-Business Management*, *16*(3), 547–578. https://doi.org/10.1007/s10257-017-0362-y
Miles, M. B., & Huberman, a M. (1994). *Qualitative Data Analysis: A Sourcebook*. *Sage Publications, Beverly Hills, California USA*. https://doi.org/10.1016/0149-7189(96)88232-2
Moe, W. W., & Schweidel, D. A. (2017). Opportunities for Innovation in Social Media Analytics. *Journal of Product Innovation Management*, *34*(5), 697–702. https://doi.org/10.1111/jpim.12405
Montoya-Weiss, M. M., & Calantone, R. J. (1994). Determinants of new product performance: A review and meta-analysis. *Journal of Product Innovation Management*, *11*(5), 397–417.
Nadav, S. (2017). Dan Ariely: People Don't Build Lasting Relationships with an Algorithm. Retrieved September 6, 2019, from https://www.linkedin.com/pulse/dan-ariely-people-dont-build-lasting-relationships-algorithm-nadav/
Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*. https://doi.org/10.1186/s40537-014-0007-7
Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance Through Text Mining. *Marketing Science*, *31*(3), 521–543. https://doi.org/10.1287/mksc.1120.0713
Nguyen, T., ZHOU, L., Spiegler, V., Ieromonachou, P., & Lin, Y. (2018). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers and Operations Research*, *98*, 254–264. https://doi.org/10.1016/j.cor.2017.07.004
Nonaka, I. (1991). The Knowledge-Creating Company. *Harvard Business Review*, *69*(6), 96–104.
Olson, E. M., Walker, O. C., & Ruekert, R. W. (1995). Organizing for Effective New Product Development: The Moderating Role of Product Innovativeness. *Journal of Marketing*, *59*(1), 48–62. https://doi.org/10.1177/002224299505900105
Oussous, A., Benjelloun, F. Z., Ait Lahcen, A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University - Computer and Information Sciences*, *30*(4), 431–448. https://doi.org/10.1016/j.jksuci.2017.06.001
Palacios Fenech, J., & Tellis, G. J. (2016). The Dive and Disruption of Successful Current Products: Measures, Global Patterns, and Predictive Model. *Journal of Product Innovation Management*, *33*(1), 53–68. https://doi.org/10.1111/jpim.12256
Pavlou, P. A., & El Sawy, O. A. (2011). Understanding the Elusive Black Box of Dynamic Capabilities. *Decision Sciences*, *42*(1), 239–273. https://doi.org/10.1111/j.1540-5915.2010.00287.x
Raisch, S., & Birkinshaw, J. (2008). Organizational ambidexterity: Antecedents, outcomes, and moderators. *Journal of Management*, *34*(3), 375–409. https://doi.org/10.1177/0149206308316058
Ransbotham, S., & Kiron, D. (2017). *Analytics as a Source of Business Innovation*. *MIT Sloan Management Review*. https://doi.org/10.1002/ece3.1283
Reid, S. E., & De Brentani, U. (2004). The fuzzy front end of new product development for discontinuous innovations: A theoretical model. *Journal of Product Innovation Management*, *21*(3), 170–184. https://doi.org/10.1111/j.0737-6782.2004.00068.x
Rindfleisch, A., O'Hern, M., & Sachdev, V. (2017). The Digital Revolution, 3D Printing, and Innovation as Data. *Journal of Product Innovation Management*, *34*(5), 681–690. https://doi.org/10.1111/jpim.12402
Ringel, M., Taylor, A., & Zablit, H. (2015). *The Most Innovative Companies 2015: Four Factors that Differentiate Leaders*. *BCG: Most Innovative Companies Series*. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=heh&AN=73929877&site=ehost-live
Robson, C. (2002). *Real World Research. 2nd. Edition. Blackwell Publishing. Malden*.
Russom, P. (2011). *The Three Vs of Big Data Analytics*.
Sambamurthy, V., Bharadwaj, A. S., & Grover, V. (2003). Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms. *MIS Quarterly*,

27(2), 237–263. https://doi.org/10.2307/30036530
Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students* (Fifth). Pearson.
Schilling, M. A. (2013). *Strategic Management of Technological Innovation* (4th ed.). New York: McGraw-Hill Irwin. https://doi.org/10.3395/reciis.v2i1.163pt
Schilling, M. A., & Hill, C. W. L. (1998). Managing the new product development process: Strategic imperatives. *Academy of Management Executive*, *12*(3), 67–81.
Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. New York, NY: Harper & Row.
Shackle, G. L. S. (1958). *Time in Economics*. Amsterdam: North-Holland.
Shackle, G. L. S. (1979). Imagination, Formalism and Choice. In *Time, uncertainty and disequilibrium*. Lexington, Mass.
Shah, R. (2019). Personal Communication - 11 July 2019. Copenhagen.
Shuradze, G., Bogodistov, Y., & Wagner, H.-T. (2018). The role of marketing-enabled data analytics capability and organisational agility for innovation - empirical evidence from german firms. *International Journal of Innovation Management*, *22*(4), 32.
Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852
Song, M., & Montoya-Weiss, M. M. (1998). Critical Development Activities for Really New versus Incremental Products. *Journal of Product Innovation Management*, *15*(2), 124–135. https://doi.org/10.1016/S0737-6782(97)00077-5.
Stremersch, S., Muller, E., & Peres, R. (2010). Does new product growth accelerate across technology generations? *Marketing Letters*, *21*(2), 103–120. https://doi.org/10.1007/s11002-009-9095-0
Tallon, P. P., & Pinsonneault, A. (2011). Competing Perspectives on the Link between Strategic Information Technology Alignment and Organisational Agility: Insights from a Mediation Model. *MIS Quarterly*, *35*(2), 463–486. Retrieved from http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=df9726e9-6d6e-4902-911b-f888c3956056%40sessionmgr4009&vid=4&hid=4109
Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* (Applied so). Thousand Oaks, CA, US: Sage Publications, Inc.
Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, *18*(7), 509–533.
Thomsen, H. (2019). Personal Communication - 23 July 2019. Copenhagen.
Trinkfass, G. (1997). *The Innovation Spiral: Launching New Products in Shorter Time Intervals*. https://doi.org/10.1007/978-3-663-09041-0
Troy, L. C., Hirunyawipada, T., & Paswan, A. K. (2008). Cross-Functional Integration and New Product Success: An Empirical Investigation of the Findings. *Journal of Marketing*, *72*(November), 132–146.
Tuarob, S., & Tucker, C. S. (2015). Automated discovery of lead users and latent product features by mining large scale social media networks. *Journal of Mechanical Design, Transactions of the ASME*, *137*(7), 1–11. https://doi.org/10.1115/1.4030049
Vas, K. (2019). Personal Communication - 3 May 2019. Copenhagen.
Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, *18*(1), 1–19. https://doi.org/10.1186/s12874-018-0594-7
Veryzer Jr., R. W. (1998). Discontinuous Innovation and the New Product Development Process. *Journal of Product Innovation Management*, *15*(4), 304–321. https://doi.org/10.1016/S0737-6782(97)00105-7
Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. fan, Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*,

356–365. https://doi.org/10.1016/j.jbusres.2016.08.009
Wedel, M., & Kannan, P. K. (2016). Marketing Analytics for Data-Rich Environments. *Journal of Marketing*, *80*(6), 97–121. https://doi.org/10.1509/jm.15.0413
Willig, C. (2001). Introducing qualitative research in psychology Adventures in theory and method. *Qualitative Research*. https://doi.org/10.1177/1468794106058877
Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, *69*(5), 1562–1566. https://doi.org/10.1016/j.jbusres.2015.10.017
Yin, R. K. (2009). *Case study research: Design and methods* (Fourth). Los Angeles: SAGE Publications.
Zhan, Y., Tan, K. H., Ji, G., Chung, L., & Tseng, M. (2017). A big data framework for facilitating product innovation processes. *Business Process Management Journal*, *23*(3), 518–536. https://doi.org/10.1108/BPMJ-11-2015-0157
Zhan, Y., Tan, K. H., Li, Y., & Tse, Y. K. (2018). Unlocking the power of big data in new product development. *Annals of Operations Research*, *270*(1–2), 577–595. https://doi.org/10.1007/s10479-016-2379-x
Zhao, J. L., Fan, S., & Hu, D. (2014). Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics*. https://doi.org/10.1080/23270012.2014.968643

Appendices

Appendix 1: Example Interview Guide Round 1

Interview Guide – Kiran Vas (Friday 3rd May)

Intro questions

- What is your role at 2021.ai?
- Could you please describe your background and previous professional experience?
  - Valcon, 2021.ai, Other
- 2021.ai company description:
  - Areas of expertise/Types of cases they typically work on

Overview: Data Analytics

Current applications of Data Analytics

- Within Data Analytics what areas have you worked on specifically?
- From your perspective, how has data analytics evolved over the past 15 years?
- What are the antecedents necessary for an organisation to implement and use data analytics?

Future applications of Data Analytics (Not necessarily linked to innovation)

- How do you see the evolution of data analytics over the medium to long term (+5 years)?

Overview: Innovation Management

- Have you undertaken any projects related to innovation?
- Could you describe the state of the art in innovation management from your perspective?

Innovation + Analytics

Current applications of data analytics + innovation

- How can data analytics be used for innovation?
  - Describe how they can be linked together
  - Type of analysis – descriptive, predictive, prescriptive etc.
- Do you have any use cases from 2021.ai or other previous experience?
- Do you see any patterns emerging in the use of data analytics for innovation?

Future applications of data analytics + innovation

- Do you see any evolving applications of data analytics for innovation in the future?
- What are the drivers behind these changes in analytics and innovation (if any)

Key Success Factors

i.e. what is needed to make the data analytics for innovation a success

- Internal (e.g. culture, capabilities and resources)
- External (industries, B2B vs B2C etc.)

Closing Questions

- Anybody else in your network that you think we should talk to?

Appendix 2: Example Interview Guide Round 2 (1/2)

Interview Guide – Hannah Haugbølle Thomsen (Tuesday 23rd July) – 60 minutes

Profile*: Head of Development and Analysis, Product Management for past 2 years. Background as a consultant at PA Consulting.*

Intro questions (5 minutes)

- What is your role at Pandora?
- Explain a bit about your background related to Data Analytics.

Section 1: BDAC for NPD (30 minutes)

RQ1: *How* can big data analytics (BDA) impact new product development (NPD) performance?

Specifically, how can big data analytics be used in the different stages of the innovation process?
In each of the stages please reflect on the type and source of the data and the method of analysis.

- How can BDA be used in the Ideation stage? Can you give concrete use cases?
  Ideation = generation and evaluation of new product ideas and further refinement of the most promising ideas into new product concepts
- How can BDA be used in the Product Development stage? Can you give concrete use cases?
  Product Development = consists of first developing the concept into a physical prototype, before testing it and utilizing feedback to arrive at the physical product
- How can BDA be used in the Product Launch stage? Can you give concrete use cases?
  Product Launch = launching the product in the market and ensuring that it is a commercial success

NPD Success Criteria:

Now looking at NPD process as a whole, how can BDAC impact this process according to the following three criteria?

1. Lower Costs (reduced product development costs)
2. Increased speed to market (less development time)
3. Product-market fit (= more sales and demand for product)

- Will the potential of BDA for NPD change within the next few years (5+)?
  E.g. through - more data, different data, technology, IoT etc.

- In what ways can BDA not impact the NPD process?

Appendix 3: Example Interview Guide Round 2 (2/2)

Section 2 - Organisational & Contextual contingency factors (20 minutes)

RQ2: *How* is the relationship between BDA and NPD performance *contingent* on organisational factors?

For each of the moderating factors outlined below, how do they influence the BDA-NPD relationship (positively or negatively) and against what criteria?

Contingency factors:

Degree of product innovativeness
Operationalisation: Consider innovativeness on a spectrum with incremental and radical at each end.
Radical → discontinuities at the industry level, BOTH introducing new technology AND creating a new market.
Incremental → only discontinuity at the level of the firm

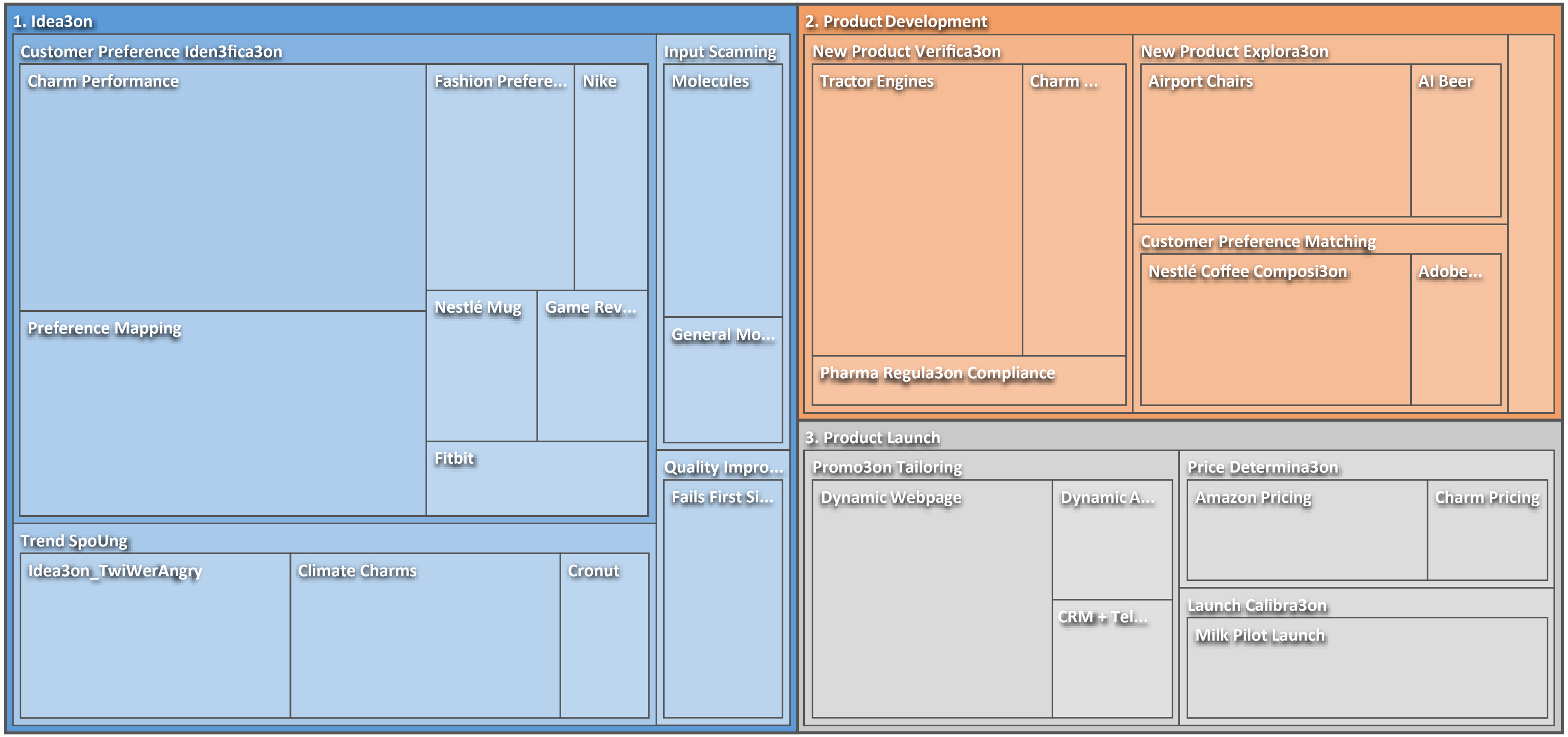
- How is the use of BDA for NPD contingent on the innovativeness of the product?
- How is the value of BDA for NPD contingent on the innovativeness of the product?

Level of Organisational agility
Operationalisation:
OA = firm's ability to detect opportunities for innovation and seize those competitive market opportunities by assembling requisite assets, knowledge, and relationships with speed and surprise.
→ How does a firm's level of Organisational agility impact the use of BDA for NPD?
I.e. very agile companies versus less agile companies.

****How important are each of these for the overall success of a firm's NPD?****

Closing Questions (5 minutes)
- Anybody else in your network (in or out of Pandora) that you think we should talk to?
- If anything comes up, could we come back to you for a second interview or generally ask more questions?

Appendix 4: Results from Nvivo Coding (Created with Nvivo)



N.B. The relative sizes of the boxes are determined by the number of coded references. For example, Ideation represents approximately half of the total area therefore roughly half of the coded references also pertain to this category.

[Page intentionally left blank]