Opportunities of Data-Driven Decision Making during the Product Innovation Process

Author: Anja Dömer
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT

Successful product innovation is a crucial element for companies to stay competitive. Nevertheless, 50-90% of new product development initiatives fail. One of the reasons is the lack of accurate information that decisions are based on during the new product development process. This study investigates by means of a systematic literature review the variety of data and data analytics that can be applied as well as the decisions that can be driven by data during the product innovation process. The results suggest that there are various opportunities for applying data-driven decision making during the product innovation process. Nevertheless, the results indicate that there is untapped potential of using unstructured data as well as applying prescriptive analytics during the product innovation process. However, the results point out that this topic still needs to be further researched, especially in the field of data-driven development and testing.

Graduation Committee members:
Matthias de Visser
Fons Wijnhoven
Liza Kayser

Keywords
Big data, data analytics, data-driven, product innovation, new product development, decision-making

1. INTRODUCTION

1.1 Problem Statement

Due to an ever-changing environment innovation has become crucial for the survival of companies (Dereli, 2015). Through innovation, companies keep their customers satisfied, gain a competitive advantage, and increase market share. In the German automobile industry in 2018, 12.8% of industry revenue was generated by new product innovations and 35.8% by imitation innovations (Statista, 2020). Consequently, innovation can be considered a key resource to generate revenue. Nevertheless, according to Heidenreich and Spieth (2013) about 50 to 90 % of product innovation projects fail due to a lack of accurate information. Because of that decision making during the innovation process is related to a high level of uncertainty.

According to Waller and Fawcett (2013) big data provides the opportunity to revolutionize decision making by making it data driven. Data-driven decision making is understood as basing decisions on the insights gained from data rather than on intuition (Provost & Fawcett, 2013). A study found out that "companies in the top third of their industry in the use of data-driven decision making were, on average, 5% more productive and 6% more profitable than their competitors" (McAfee and Brynjolfsson, 2012, p. 6).

Even though the potential benefits of leveraging big data for decision making during the product innovation process are significant, research has not yet provided a comprehensive systematic literature review on data-driven decision making during the product innovation process.

1.2 Research Objectives & Question

The goal of this research is to present how data can drive decision making during the product innovation process. The outcome of this research will be an overview of the current literature about business analytics that can be applied to decisions made during the new product development process. Thereby, a focus is set on the variety of data and the type of analytics that can be applied. During this research, the following research question will be answered:

How can (big) data drive decision making during the product innovation process?

To answer this question the following sub-questions have been formulated:

- Which product innovation decisions can be driven by data?
- Which variety of data can be used to drive product innovation decisions?
- What kind of analytics can be used to drive product innovation decisions?

1.3 Research Relevance

1.3.1 Academical Relevance

This research aims to present current literature about (big) data and (big) data analytics, and the application to product innovation decisions. Thereby, it aims to identify current gaps in literature.

1.3.2 Practical Relevance

This research aims to provide examples for companies on how (big) data can be applied in the context of product innovation. It intends to serve as an orientation for companies by providing insights on potential (big) data sources and (big) data analytics that can be used during the product innovation process to drive decision making.

1.4 Outline of the Paper

The following paper is structured into seven main sections. Section 2 *theoretical background* elaborates on literature related to the fields of innovation and data to provide the reader with a broad understanding of these topics. Section 3 *research methodology* explains the applied research methods in this study. Section 4 *results* elaborates on the literature selection process and analyses the selected literature. Section 5 *discussion* reviews the literature results while section 6 *conclusion* concludes this study. Section 7 refers to limitations as well as possible future research topics.

2. THEORETICAL BACKGROUND

2.1 Innovation

The word "innovation" comes from the Latin word novus which can be translated to "new" (Gopalakrishnan, 1994). In the organizational context an innovation can be a new product, a new process of production, a new material in an existing product, a new workflow, or a new method (Kline, 2010).

Innovation literature commonly emphasizes the difference between an invention and an innovation. While an invention is the first discovery of an idea for a new product or process, innovation is the first application of an invention in practice (Fagerberg, 2006). Fagerberg claims that "to be able to turn an invention into an innovation, a firm normally needs to combine several different types of knowledge, capabilities, skills, and resources" (Fagerberg, 2006, p.4). Paul Trott claims that innovation needs to be viewed as a process that needs to be managed rather than as solely the outcome of a process (Trott, 2012). Therefore, he defines innovation as "the management of all activities involved in the process of idea generation, technology development, manufacturing, and marketing of a new (or improved) product or manufacturing process" (Trott, 2012, p.15). Within an organization there are various kind of innovations such as product, process, or business model innovation. This research will focus on product innovation.

2.1.1 Product Innovation

Paul Trott defines product innovation as the development of a new or improved product (Trott, 2012). An enhanced product is referred to as a product innovation if the product's characteristics or intended uses are enhanced. This comprises significant enhancements in technical product specifications or functional product characteristics (Anderson, 2014). The output of a product innovation can be a good or a service (Schilling, 2019). Due to significant differences in product characteristics of goods and services such as intangibility, heterogeneity, perishability, and consumption (Trott, 2012) extant literature differentiates between the development process of a good and the development process of a service. This research will focus on the development process of goods. In the following the terms product and good are used interchangeably.

2.1.2 The Process of Product Innovation

The development process of a product can be described as "the process of transforming business opportunities into tangible products" (Trott, 2012, p. 418). A widely used approach to describe the new product development process of products is the "stage-gate-process" from Cooper that has been published for the first time in 1988. The typical stage-gate process divides new product development into five stages and five gates starting with the discovery of opportunities for product innovations and ending with the post-launch review of the innovation (Cooper, 1988).

The discovery stage, also called ideation, is focused on the discovery of new market opportunities and therefore, the rough development of ideas for innovative products. The Scoping or Concept stage sharpens the product concept and develops specific product requirements and designs. The Business Case stage deals amongst others with the financial and profitability analysis of the product. The Development stage develops the prototype of the new product while the Testing stage tests the prototype and validates it against customer requirements. The Launch stage is focused on bringing the new product to the market (Cooper, 1988).

For a product to move to the next stage in the development process it needs to pass the associated gate. Each gate refers to specific deliverables that the innovation team needs to provide as well as criteria that a product needs to fulfill, otherwise the innovation project might either be put on hold or cancelled completely (Cooper, 1988). The idea is to recognize already at an early stage if an innovation initiative is likely to fail to be able to cancel the project early and therefore, safe time, resources, and costs. Up until now, Cooper has developed updated models of the stage-gate-process that incorporate new emerging working methods such as agile development (Cooper & Sommer, 2018). Moreover, in 2008 he represented three variations of the stage-gate-process according to the nature of the new product projects: Stage-Gate (Full), Stage-Gate (Xpress), and Stage-Gate (Lite) (Cooper, 2008).

Next to the stage-gate process, extant literature provides various models and concepts for new product development. The models are supposed to serve as guidelines for companies during the innovation process. They differentiate themselves by focusing on diverse aspects of the process or representing other viewpoints. Further NPD models include, but are not limited to market-pull or technology push-model (focused on the driver of innovation), network-model (focused on the knowledge creation), departmental-model (focused on the departments involved) or activity-model (focused on the innovation activities) (Trott, 2012).

Since the stage-gate process has been published first it received criticism from multiple authors. Amongst others, authors criticized that the process is a rigid and linear process that companies are following step by step. They describe it as inflexible, non-adaptable and outdated (Cooper, 2008). However, in his paper from 2008 Cooper reacts to these authors and explains that the stage-gate process has evolved over the years. He refers to the process as a map that guides the way rather than a lock-step process that strictly needs to be followed (Cooper, 2008). Cooper (2008) emphasizes that companies need to tailor the process according to their requirements which can also mean skipping stages or rearranging the order of them. Furthermore, he points out that even though the process is visualized linear, activities can take place parallel as well as stages can overlap and iterate (Cooper, 2008).

Even though many new processes and tools have been developed to describe and guide the product innovation process the stage-gate process seems to be the most dominant and resilient system in the literature. Therefore, this thesis focuses on the stage-gate model as an innovation process model.

2.1.2.1 Decisions During the Innovation Process

Even though the innovation processes might vary between companies as well as within a company over time, the decisions that need to be made during the innovation process of products are usually focused on similar issues (Krishnan & Ulrich, 2001). Krishnan and Ulrich (2001) conducted a literature review on product and project related decisions during the innovation process and identified dominant decisions reoccurring in literature.

Krishnan and Ulrich (2001) differentiate between two categories of decisions. The first category includes decisions that are associated with the actual development of the product and the second category refers to decisions related to the planning of the development project.

The first category can further be divided into five areas:

1. Concept development
2. Supply chain design
3. Product design
4. Performance testing and validation
5. Production ramp-up and launch

Concept development decisions refer to decisions that determine product requirements and the product’s basic physical design. Supply chain decisions refer to decisions about the acquisition and the flows of materials, the purchase of services and intellectual property from third parties, the selection of suppliers as well as the general design of a production and distribution system. Product design decisions refer to decisions that determine design parameters of the entire product as well as the components of its assembly. Performance testing and validation decisions refer to decisions about how to prototype the product to test it. Product launch and production ramp-up decisions refer to decisions about test marketing and the introduction of the product into a new market (Krishnan & Ulrich, 2001).

The second category, project development decisions, can be further divided into three areas:

1. Product strategy planning
2. Product development organization
3. Project management

Product strategy and planning decisions refer to decisions that set the strategic direction for the development project to ensure alignment with the strategy of the company, i.e. the decision about the right market. Product development organization decisions refer to decisions about the context in which the project will be executed such as the project team. Project management decisions refer to structural project decisions such as milestones, team communication or project monitoring (Krishnan & Ulrich, 2001). Figure 1 shows the attribution of the innovation decision categories by Krishnan and Ulrich (2001) to the stages of the stage-gate process by Cooper (2008).

It can be said that with each decision made during the product innovation process the project evolves and moves to another stage of the process. Each stage means more allocation of time and resources into the project than in the proceeding stage and therefore, more costs (Cooper, 2008). Product innovation teams need to make fast decisions to keep the duration of the product innovation process as short as possible and they need to make the right decisions to keep the risk of project failure as low as possible. “Data are widely considered to be a driver of better decision making” (Waller & Fawcett, 2013, p. 77) and therefore, provide the opportunity to achieve fast and accurate decision making.

2.2 Data

Companies that make decisions based on data are also referred to as data-driven organizations. Data-driven organizations are more productive and profitable, can better adapt to changes in the environment such as threats and opportunities, and can better forecast the future (Schneider & Valacich, 2018, p. 254). Factors that might not only benefit the day-to-day business but also the new product development process. To understand how data can positively affect decision making the following sections define data as well as data analytics.

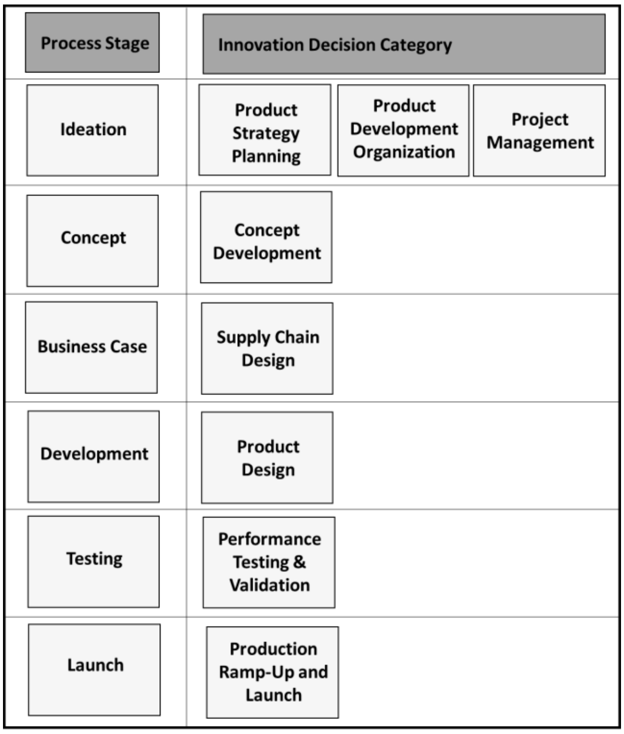

Figure 1: Own attribution of process stages by Cooper (2008) and innovation decision categories by Krishnan & Ulrich (2001)

2.2.1 (Big) Data Definition

According to the Merriam-Webster dictionary (n.d.) data can be defined as "factual information (such as measurements or statistics) used as a basis for reasoning, discussion or calculation" and as "information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful". This definition emphasizes that without transforming raw data via analysis into meaningful insights, data is of no or low value to a company.

Per day approximately 2.5 Exabytes of data are generated (IBM, 2016). Due to the tremendous amount of data literature also refers to this data as "big data". Big data is described as "the next big thing in innovation" (Gobble, 2013) and experts claim that it will lead to "a revolution in science and technology" (Ann Keller et al., 2012).

The term "big data" has first been defined by Laney in 2001 via the 3Vs: volume, velocity, and variety.

Volume refers to the size of the stored data (Wamba et al., 2015). Commonly big data encompasses multiple terabytes and petabytes. However, there are no clear criteria that determine which data size differentiates data from big data.

Velocity refers to the speed of the generation of data and the necessary speed to analyze and act upon it (Wamba et al., 2015). New technologies such as the Internet of Things lead to continuous and unprecedented data creation. In general, it can be said that big data is generated at a high velocity.

Variety describes "the structural heterogeneity in a dataset" (Gandomi, 2015, p. 138). Humans may have created the data via i.e. survey or logs or machines via i.e. weather data or road vision (Shah, 2020). The variety ranges from structured over semi-structured to unstructured data whereas structured data is considered data in an organized form such as data in spreadsheets and unstructured data is considered data in an unorganized form such as text, images, audio and video (Gandomi, 2015). Semi-structured data is a combination of both. Unstructured data encompasses 95% of all existing data (Miah et al, 2017). While structured data usually is collected via devices that are connected to the Internet of Things such as scanners, sensors, GPS tracker etc., unstructured data lies hidden in data that is mostly extracted from social media such as social networks or blogs (Qi et al., 2016). Big data is referred to have a wide variety of different data sources.

Further dimensions that can be used to describe big data are veracity, variability (and complexity), and value.

Veracity refers to the unreliable character of some data sources because they are based on human judgement such as the opinion of customers (Gandomi, 2015).

Variability describes the varying rates at which data is generated whereas *complexity* refers to the difficulty of combining and analyzing various data sources (Gandomi, 2015).

Value refers to the worth of (big) data. It can be said that (big) data in its raw form has a low value. However, value is created by analyzing it to gain new insights and support decision making (Iqbal et al., 2020).

Even though big data can be described based on these different dimensions, there have been no clear thresholds that clearly separate data from big data. Literature even refers to the definition as a "moving definition" as everyday more data is generated and what is considered big data today might only be regarded as data in the future. Consequently, the definition changes over time and based on the applied industry (Sheng et al., 2017). Therefore, the research of this thesis does not differentiate between data and big data.

2.2.2 Data Analytics

To derive value from (big) data, a company needs to use various techniques that transform (big) data into valuable insights. Literature refers to these as Data Analytics. Data Analytics are "the use of data, information technology, statistical analysis, quantitative methods and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions" (Evans, 2017, p. 3). Wamba et al. defines Big Data Analytics as "a holistic approach to managing, processing, and analyzing the 5V data-related dimensions (i.e. volume, variety, velocity, veracity and value) to create actionable ideas for delivering sustained value, measuring performance and establishing competitive advantages" (Wamba et al., 2017, p. 356). Kwon et al. defines it as "technologies (e.g. database and data mining tools) and techniques (e.g. analytical methods) that a company can employ to analyze large scale, complex data for various applications intended to augment firm performance in various dimensions" (Kwon, 2014, p. 387). Literature sometimes separates Data Analytics and Big Data Analytics due to the size of the analyzed data and the increasing challenges for the analytics to cope with big data. Nevertheless, both aim to gain insights from data via the analysis of it. Due to the fuzzy thresholds separating data and big data this research does not differentiate between Data Analytics and Big Data Analytics.

(Big) Data Analytics can be categorized according to the data that is analyzed or the general purpose of the Analytics.

2.2.2.1 Data Analytics According to the Analyzed Data

Analytics categorized according to the data that is analyzed include, but are not limited to text analytics, audio analytics, video analytics and social media analytics.

Text analytics aim to gain insights from unstructured textual data such as blogs, e-mails, survey responses or corporate documents (Gandomi, 2015). *Audio analytics* aim to gain insights from unstructured audio data such as recorded customer calls from call centers while *video analytics* aim to gain insights from recorded videos such as videos from the surveillance cameras in a store (Gandomi, 2015). *Social media analytics* aim to gain insights from structured and unstructured data from online platforms such as social networks, blogs, media sharing or review sites. (Stieglitz et al., 2018).

2.2.2.2 Data Analytics According to the General Purpose of the Analytics

Analytics categorized according to the general purpose of the analytics include, but are not limited to descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics.

Descriptive analytics "categorize, characterize, consolidate, and classify" past and current data to provide meaningful information about the business performance (Evans, 2017, p. 8). They aim to identify i.e. opportunities, threats, or trends.

Predictive analytics analyze historical and current data to predict the future. The analytics use the outcomes of the analysis of the past data and then apply it to a situation in the future (Shah, 2020; Evans, 2017). Predictive analytics can be used for i.e. predicting consumer behavior, demand, or future costs.

Diagnostic analytics, also known as causal analytics, aim to identify why a specific outcome occurred (Shah, 2020), for example to analyze why did customers react in a certain way to a new product.

Prescriptive analytics aim to identify the optimal outcome of a given situation to determine the best course of action (Shah, 2020). The analytics can be used to calculate i.e. the best pricing strategy (Evans, 2017).

The current literature that has recognized the potential of data and data analytics during the innovation process focuses for example on guidelines for facilitating the use of big data during the product innovation process (Tan & Zhan, 2017), data analysis methods to develop services (Meierhofer & Meier, 2019), business model innovation via data (Bhimani et al., 2019), the indirect impact of business analytics on companies' competitive advantage (Duan et al., 2020) or the impact of big data on companies' innovation performance (Niebel et al., 2019). Nevertheless, there has not been executed a systematic literature review on specific application opportunities for data-driven decision making alongside the entire product innovation process that also considers application opportunities for different types of data varieties and analytics. Therefore, this research has been executed.

The following section describes the research methods used during this study.

3. RESEARCH METHODOLOGY

To identify and analyze relevant research on data-driven decision making during the product innovation process, a systematic literature review has been conducted. Therefore, the approach towards a literature review of Brocke et al. (2009) has been used to structure the review in this research. This approach includes the following steps: definition of review scope, conceptualization of topic, literature search, literature analysis and synthesis, as well as research agenda.

3.1 Definition of review scope

According to Brocke et al. (2009) the scope of a review can be best defined via the following characteristics: focus, goal, organization, perspective, audience, and coverage.

The *focus* of this literature review will be on the research outcomes of identified literature. The research outcomes of extant literature combined will help to answer the question of how data-driven decision making can drive the product innovation process.

The *goal* of the literature review is to integrate and synthesize literature to provide an overview of extant literature in the field of data-driven decision making during the product innovation process as well as to identify current research gaps. Moreover, this literature review aims to present central outcomes related to data-driven decision making.

The *organization* of the literature review will be conceptual. The concepts are the decisions that must be made during the innovation process according to the literature review of Krishnan and Ulrich (2001).

The *perspective* of this literature review is to identify and analyze literature that supports the idea of data-driven decision making during the innovation process.

The *audience* of this literature review is expected to be practitioners and general scholars.

The *coverage* is representative which means the literature review intends to represent the most relevant literature in the field of data-driven decision making during the product innovation process.

3.2 Conceptualization of the topic

According to Brocke et al. (2009), before beginning with the literature search, key terms should be defined and conceptualized to gain a broad understanding of the topic. The key concepts in this research are data and innovation. An initial search has been done in Scopus.

To develop a better understanding of innovation itself, the process as well as the decisions during the process, the following keywords have been used: "innovation", "innovation process model", "innovation management", "new product development" and "new product development and activities". To ensure a comprehensive overview, there has been no limitation in years, solely a limitation to the subject field "Business, Management and Accounting". Moreover, only literature that had been cited more than 40 times was reviewed and considered in this review. The threshold of 40 times had been set to select papers that have been cited a significant amount of times by other authors and therefore, indicate valid results. The threshold was not set lower as it would have been out of the scope of this study to review this amount of papers. Literature published after 2018 had been reviewed no matter how many times it had been cited due to its recent publication. If literature has met these requirements, it has been selected if it contributed to the understanding of innovation, product innovation and product innovation process models.

To develop a better understanding of the field of data in the context of organizations the following keywords have been used: "data science", "data", "big data", "big data analytics" and "data-driven decision making". As the technologies supporting data analytics but also data itself is continuously evolving due to frequent innovations, the literature search has been limited to the years 2016 to 2020 to avoid investigating outdated literature. Moreover, literature has been limited to the subject field "Business, Management and Accounting". Due to the recency of publication, there are no requirements for a minimum amount of citations. If all previously mentioned requirements have been met, literature has been selected based on title and abstract.

The outcome of the initial research is presented in section 2 *related literature*. To visualize the relevant concepts for this research concept mapping has been used as can be seen in figure 2.

3.3 Literature Search

To identify relevant literature for this review, literature has been searched in Scopus which claims to deliver "the most comprehensive overview of the world's research output" amongst others in the field of science and technology (Elsevier, 2020). Therefore, it includes all relevant scientific journals and books and makes a separate journal search redundant.
The literature search has been separated based on the decision categories that have been identified in the literature review of Krishnan and Ulrich (2001). The decision categories have been grouped according to the field of expertise they are relating to, i.e. Concept Development and Launch (Marketing & Sales), Supply Chain Design & Production Ramp-Up (Logistics & Operations), Product Design & Performance Testing and Validation (Technology Development), Product Development Organization (Human Resources), Product Strategy and Planning & Project Management (Strategy Building). The keywords for the systematic literature search consist of various descriptions of these fields as well as a keyword relating to the field of data. The literature search has been limited to the subject field "Business, Management and Accounting" as well as to the years 2016 to 2020. If all requirements have been met, it has been assessed whether the literature can be related to one of the decisions identified by Krishnan and Ulrich. This assessment was done first via title, then via abstract and then via the full text. The selected literature includes empirical as well as non-empirical literature. There has not been a focus on a specific kind of research method applied in the literature. To ensure that literature that might still be relevant but has been published before 2016 and therefore had been excluded from the database search, is not ignored, a forward and backward search has been done. That means if relevant research outcomes have been mentioned in the identified literature, the articles have been looked up and checked. If they have been assessed as relevant, they have been included in this literature review.

The search strings used for the analysis are represented in the following:

Concept Development, Launch
1. "marketing" and "big data analytics"
2. "marketing" and "data-driven"
3. "marketing analytics"
4. "customer analytics"

Supply Chain Design, Production Ramp-Up
1. "supply chain" or "operations" and "big data analytics"
2. "supply chain" and "data-driven"
3. "production" and "data-driven"
4. "operations analytics" or "production analytics" or "supply chain analytics"

Product Design, Performance Testing and Validation
1. "engineering" and "big data analytics"
2. "engineering" or "prototyping" or "product design" or "development" and "data-driven"
3. "engineering analytics" or "research analytics" or "design analytics"

Product Development Organization Decisions
1. "human resource" and "big data analytics"
2. "human resource" and "data-driven"
3. "human resource analytics"
4. "hr analytics" or "people analytics" or "workforce analytics" or "talent analytics"

Product Strategy and Planning, Project Management
1. "project" or "strategy" and "big data analytics"
2. "project" or "strategy" and "data-driven"
3. "project analytics" or "strategy analytics"
4. "project" or "strategy" and "business intelligence"

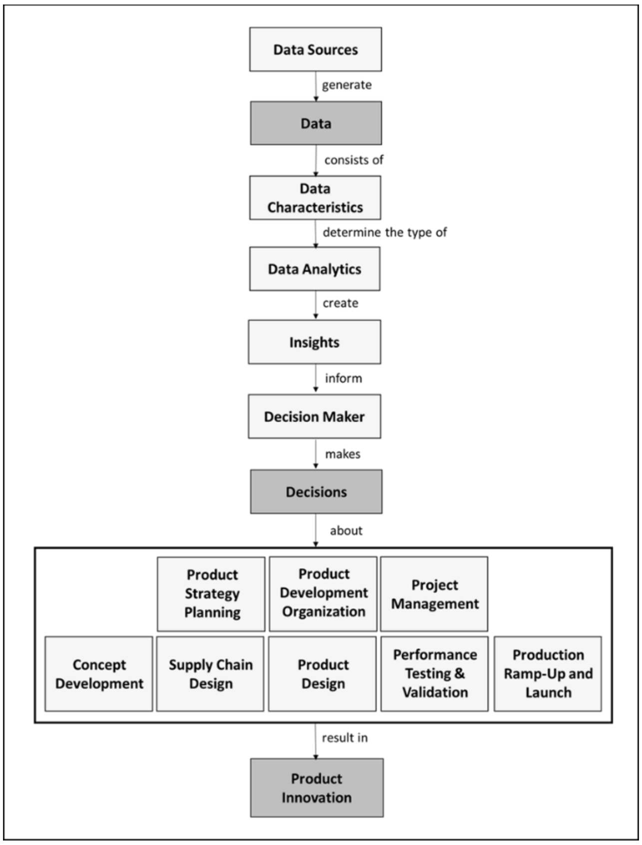

Figure 2: Concept map "data, product innovation decisions"

3.4 Literature Analysis

The aim of the literature analysis is to answer the sub-questions of the main research question that are mentioned in section 1: Which product innovation decisions can be driven by data? Which variety of data can be used to drive product innovation decisions? What kind of analytics can be used to drive product innovation decisions?
To answer these questions with the selected literature, a concept matrix has been created for each search blog. The concept matrices contain the year of publication, author, the type of analytics used to drive the decisions (descriptive, diagnostic, predictive, prescriptive) and the variety of the analyzed data (unstructured, structured). Semi-structured data has not been considered in the concept matrices as semi-structured data has not been clearly identifiable in literature.

4. RESULTS

This section is left blank intentionally.

5. DISCUSSION

The literature analysis shows that decisions can be driven by data during the ideation stage (partly), the concept development stage, the business case stage, and the launch stage. During the literature selection process, no papers were identified that would describe data-driven decision making of product strategy and planning decisions (ideation stage) as described by Krishnan and Ulrich (2001). These decisions would be amongst others driven by data about the market, about customers and about internal operations which are proven to support decision making in further stages of the process. For example, the concept development stage is driven by data about customers to sharpen the product idea. However, the same data insights can also be used in the ideation stage to discover market opportunities even though literature currently has not investigated a connection between these data insights and decisions yet. Another decision category for which no literature could be identified, and which belongs to the ideation stage is "project management". These decisions might also be driven by data about internal operations which are proven to drive decisions during the business case stage. Furthermore, no papers were dedicated to describing and proving data-driven decision making during the development stage and the testing stage. Nevertheless, this does not mean that data cannot drive the decisions in these stages. The literature review based on the applied methodology in this study indicates that data-driven development and data-driven testing is an unexplored research field. However, when having a closer look at the decisions that have been identified by Krishnan & Ulrich (2001) for these two stages it can be assumed that these decisions can also be driven via i.e. data about customers, business operations and external parties. Consequently, the correct question to ask might not be which decisions can be driven by data and analytics but for which decisions during the innovation process does it make sense to be data-driven. This research highlighted various opportunities for data to make decision making during the innovation process more fact-based and real-time. Nevertheless, there are also various challenges that come along with data-driven decision making. One of the main challenges is its implementation during the innovation process as companies need to have the capabilities for it. The capabilities include i.e. the technological infrastructure within a company to collect, store and analyze the data (Akter et al., 2016) or the innovators' data literacy skills to understand, interpret and use data (Janssen et al., 2017).

Having a look at the entire product innovation process a slight tendency towards analyzing structured data can be identified. As mentioned in section 2, 95% of the existing data in the world is unstructured which means by collecting and analyzing tendentially rather structured data companies do not reach full potential by far. On the one hand, unstructured data is another potential source to gain valuable insights that are available in a huge amount to companies such as on social media or in corporate documents The large amount of unstructured data that is publicly available for use might seem as an attractive option. On the other hand, it is questionable whether unstructured data provides more or as valuable insights as structured data. As unstructured data is more complex to analyze and requires individual analysis methods it might be more difficult and costly to gain insights from it. Especially during the product innovation process this can be an important factor. The potential insights of unstructured data are used to drive decisions for a specific product innovation and might not be used for any other purpose. A company needs to ensure that the value gained from unstructured data is higher than the investment that needs to be made to gain insights from these. Again, it is not a question about which data can be used during the product innovation process, but which data (insights) provide more value for a company.

Also, as already mentioned in section 4.6, the analytics that are used during the product innovation process are either descriptive, diagnostic, or predictive. The literature review indicates that there is almost no application of prescriptive analytics during the product innovation process yet. Nevertheless, there might be potential for applying prescriptive analytics. For example, during the ideation stage and the concept development stage prescriptive analytics might show future scenarios of concepts and indicate the best scenario to develop further. During the business case stage prescriptive analytics have the potential to develop various scenarios related to production matters or financial decisions while during the testing stage prescriptive analytics might be valuable to identify the most optimal version of a prototype. During the launch stage prescriptive analytics can indicate strategies related to the most optimal time and location to enter a market with a new product. Prescriptive analytics are valuable during the product innovation process as they create scenarios and compare them with each other. Thereby, these analytics can help the decision maker to choose the most optimal scenario. One potential reason why prescriptive analytics are not applied yet could be that they are more complex to implement than descriptive, diagnostic, or predictive analytics. In general, it can be said that any type of analytics can be applied in any innovation stage. Nevertheless, applying specific analytics in certain stages might make more sense than in others. Here as well, it is not a question of which data analytics can be used but which analytics provide more benefits for a company.

6. CONCLUSION

The aim of this research is to answer the research question *"How can (big) data drive decision making during the product innovation process?"*. The background research at the beginning of this thesis showed that there is a variety of data that can provide insights for decision makers by being analyzed via different type of analytics. Therefore, a systematic literature review has been done to identify the variety of data, the type of analytics and the decisions that can be driven by data during the product innovation process. The systematic literature review has been structured according to the decision categories during an innovation process identified by Krishnan and Ulrich (2001).

As discussed in the previous section, there are various opportunities for data-driven decision making during the product innovation process. Two results of the literature review that especially stood out were the slight tendency of using structured data during the product innovation process as well as the untapped potential of applying prescriptive analytics. The literature review presented various ways of how a variety of data and data analytics can be applied to innovation decisions to drive the product innovation process. Nevertheless, this study also revealed that the topic still needs to be further researched. Even though there are various opportunities for applying data-driven decision making the exact value of data-driven decisions, of unstructured or structured data and of specific analytics during the product innovation process is not known yet.

7. LIMITATIONS & FUTURE RESEARCH

7.1 Limitations

Firstly, as already mentioned in the previous sections the scope of this study is limited. This research only indicates various application opportunities for data-driven decision making during the product innovation process. Still, it does not investigate the link between the variety of data, the type of analytics and the value it has for each specific decision. Secondly, this research is only based on literature. The successful application of data-driven decision making during the innovation process is still questionable and requires further research. Thirdly, the systematic literature review has been structured and executed based on the innovation process decisions that have been identified by Krishnan and Ulrich 19 years ago in 2001.

Therefore, it is questionable if the paper still reflects all dominant decisions that need to be made during the innovation process. Moreover, as some decision categories (i.e. concept development and launch) needed similar insights literature has been searched and analyzed for both at once even though the decision categories were assigned to different product innovation stages. This makes it difficult to provide a clear statement about the applied analytics and data variety in each product innovation stage. Fourthly, the systematic literature review is limited by the selected keywords and filters. In the initial search the amount of selected papers is quite high, however after applying the filters of year and subject area the amount of papers decreases tremendously. Even though a backward and forward search has been done, there is a possibility of having missed relevant literature. Nevertheless, broadening up the keywords and filters would have been out of the scope of this research.

7.2 Future Research

As has been identified in the literature selection process, there is no literature specifically referring to the product design decisions (development stage), performance testing and validation decisions (testing stage), product strategy and planning as well as project management decisions (ideation stage). To fully discover the potential of data-driven decision making during the new product development process research needs to be done in these stages. Moreover, research is needed that investigates the value of the variety of data and the type of analytics for specific decisions in each stage. Next to that, it needs to be investigated whether the value of data-driven decision making during the product innovation process varies according to different industries, company sizes or even other product innovation process models and working methods such as agile product development.

8. ACKNOWLEDGEMENTS

I would like to finish this thesis by acknowledging the people that have supported me throughout the process. Firstly, I would like to thank my first supervisor Matthias de Visser for providing me with the freedom to choose my research focus as well as for his extensive efforts in guiding me to define that focus. Secondly, I would like to thank my third supervisor Liza Kayser for providing valuable insights from her own research experience and her ongoing enthusiasm to make this thesis the best it could be. I would like to express my sincere gratitude to both for their patience during these somewhat turbulent times, for repetitively providing constructive feedback and for their willingness to answer all my questions at any time. Thirdly, I would like to thank my second supervisor Fons Wijnhoven for providing me with final feedback at the end of this thesis to finalize this work.

Lastly, I would like to thank my family and friends for their continuous support, their genuine love and their encouragement. Without them I would not have been able to finalize this thesis.

9. REFERENCES

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

Anderson, N., Potočnik, K., & Zhou, J. (2014). Innovation and creativity in organizations: A state-of-the-science review, prospective commentary, and guiding framework. Journal of management, 40(5), 1297-1333.

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: why HR is set to fail the big data challenge. Human Resource Management Journal, 26(1), 1-11.

Ann Keller, S., Koonin, S.E., Shipp, S., (2012). Big data and city living–what can it do for us? Significance 9 (4), 4–7**.**

Arco, M. D., Presti, L. L., Marino, V., & Resciniti, R. (2019). Embracing AI and Big Data in customer journey mapping: From literature review to a theoretical framework.

Bhimani, H., Mention, A. L., & Barlatier, P. J. (2019). Social media and innovation: A systematic literature review and future research directions. Technological Forecasting and Social Change, 144, 251-269.

Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2020). Supply chain data analytics for predicting supplier disruptions: A case study in complex asset manufacturing. International Journal of Production Research, 58(11), 3330-3341.

Brocke, J. V., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process.

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. Production and Operations Management, 27(10), 1868-1883.

Cao, G., Duan, Y., & El Banna, A. (2019). A dynamic capability view of marketing analytics: Evidence from UK firms. Industrial Marketing Management, 76, 72-83.

Cooper, R. G. (1988). The new product process: a decision guide for management. Journal of Marketing Management, 3(3), 238-255.

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(3), 213-232.

Danaher, B., Huang, Y., Smith, M. D., & Telang, R. (2014). An empirical analysis of digital music bundling strategies. Management Science, 60(6), 1413-1433.

Deloitte (2015). Global Human Capital Trends.

Dereli, D. D. (2015). Innovation management in global competition and competitive advantage. Procedia-Social and behavioral sciences, 195, 1365-1370.

Dahlbom, P., Siikanen, N., Sajasalo, P., & Jarvenpää, M. (2019). Big data and HR analytics in the digital era. Baltic Journal of Management.

Duan, Y., Cao, G., & Edwards, J. S. (2020). Understanding the impact of business analytics on innovation. European Journal of Operational Research, 281(3), 673-686.

Elsevier (2020). Scopus: Content Coverage Guide.

Evans, J. R. (2017). Business analytics. Pearson.

Fagerberg, J. (2006). Innovation: A Guide to the Literature. The Oxford Handbook of Innovation, p. 1-29.

Fernandez, V., & Gallardo-Gallardo, E. (2020). Tackling the HR digitalization challenge: key factors and barriers to HR analytics adoption. Competitiveness Review: An International Business Journal.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International journal of information management, 35(2), 137-144.

Gobble, M.M, (2013). Big data: the next big thing in innovation. Res. Technol. Manag. 56 (1), 64–66

Gopalakrishnan, S., & Damanpour, F. (1994). Patterns of generation and adoption of innovation in organizations: Contingency models of innovation attributes. Journal of engineering and technology management, 11(2), 95-116.

Hamilton, R. H., & Sodeman, W. A. (2020). The questions we ask: Opportunities and challenges for using big data analytics to strategically manage human capital resources. Business Horizons, 63(1), 85-95.

Heidenreich, S., and Spieth, P. (2013). Why innovations fail-the case of passive and active innovation resistance. International Journal of Innovation Management17(05): 226–336.

Hicks, C. (2018). Predicting knowledge workers' participation in voluntary learning with employee characteristics and online learning tools. Journal of Workplace Learning.

IBM, (2016). Big Data and Analytics. Retrieved from http://www-01.ibm.com/software/data/bigdata/ what-is-big-data.htm

Iqbal, R., Doctor, F., More, B., Mahmud, S., & Yousuf, U. (2020). Big data analytics: Computational intelligence techniques and application areas. Technological Forecasting and Social Change, 153, 119253.

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. Journal of Business Research, 70, 338-345.

Johnson, D. S., Muzellec, L., Sihi, D., & Zahay, D. (2019). The marketing organization's journey to become data-driven. Journal of Research in Interactive Marketing.

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. International Journal of Operations & Production Management.

Kakatkar, C., & Spann, M. (2019). Marketing analytics using anonymized and fragmented tracking data. International Journal of Research in Marketing, 36(1), 117-136.

Kline, S. J., & Rosenberg, N. (2010). An overview of innovation. In Studies On Science And The Innovation Process: Selected Works of Nathan Rosenberg, 173-203.

Krishnan, V., & Ulrich, K. T. (2001). Product development decisions: A review of the literature. Management science, 47(1), 1-21.

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.

Laney D. (2001). 3D data management: Controlling data volume, velocity and variety. In: Meta Group. Retrieved from http://blogs.gartner.com/douglaney/files/2012/01/ad949-3D-Data-ManagementControlling-Data-Volume-Velocity-and-Variety.pdf

Leung, K. H., Mo, D. Y., Ho, G. T. S., Wu, C. H., & Huang, G. Q. (2020). Modelling near-real-time order arrival demand in e-commerce context: a machine learning predictive methodology. Industrial Management & Data Systems.

Liu, C., AbouRizk, S., Morley, D., & Lei, Z. (2020). Data-Driven Simulation-Based Analytics for Heavy Equipment Life-Cycle Cost. Journal of Construction Engineering and Management, 146(5), 04020038.

Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. Tourism Management, 71, 337-347.

Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR Analytics. The International Journal of Human Resource Management, 28(1), 3-26.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. Harvard business review, 90(10), 60-68.

McColl-Kennedy, J. R., Zaki, M., Lemon, K. N., Urmetzer, F., & Neely, A. (2019). Gaining customer experience insights that matter. Journal of Service Research, 22(1), 8-26.

Meierhofer, J., & Meier, K. (2017). From data science to value creation. International Conference on Exploring Services Science. Springer, Cham, 173-181

Merriam Webster Dictionary (n.d.). Data. Retrieved November 15, 2020, from https://www.merriam-webster.com/dictionary/data

Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A big data analytics method for tourist behaviour analysis. Information & Management, 54(6), 771-785.

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.

Niebel, T., Rasel, F., & Viete, S. (2019). BIG data–BIG gains? Understanding the link between big data analytics and innovation. Economics of Innovation and New Technology, 28(3), 296-316.

Qi, J., Zhang, Z., Jeon, S., & Zhou, Y. (2016). Mining customer requirements from online reviews: A product improvement perspective. Information & Management, 53(8), 951-963.

Pessach, D., Singer, G., Avrahami, D., Ben-Gal, H. C., Shmueli, E., & Ben-Gal, I. (2020). Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. Decision Support Systems, 113290.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big data, 1(1), 51-59.

Schilling, M. A., & Shankar, R. (2019). Strategic management of technological innovation. McGraw-Hill Education.

Schneider, C., & Valacich, J. S. (2018). Information systems today: managing in the digital world (Vol. 8). Pearson Education.

Shah, C. (2020). A Hands-On Introduction to Data Science. Cambridge University Press.

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. International Journal of Production Economics, 191, 97-112.

Sivathanu, B., & Pillai, R. (2019). Technology and talent analytics for talent management–a game changer for organizational performance. International Journal of Organizational Analysis.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. International journal of information management, 39, 156-168.

Statista (2020). Share of new products and imitation innovations in the total revenue of the automobile production industry in Germany from 2008 to 2018, Retrieved from https://www.statista.com/statistics/589548/new-products-share-car-industry-revenue/

Tan, K. H., & Zhan, Y. (2017). Improving new product development using big data: a case study of an electronics company. R&D Management, 47(4), 570-582.

Trott, P. (2012). Innovation management and new product development (5th ed.). Financial Times/Prentice Hall.

Trabucchi, D., Buganza, T., & Pellizzoni, E. (2017). Give Away Your Digital Services: Leveraging Big Data to Capture Value New models that capture the value embedded in the data generated by digital services may make it viable for companies to offer those services for free. Research-Technology Management, 60(2), 43-52.

Trusov, M., Ma, L., & Jamal, Z. (2016). Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting. Marketing Science, 35(3), 405-426.

Van den Heuvel, S., & Bondarouk, T. (2017). The rise (and fall?) of HR analytics. Journal of Organizational Effectiveness: People and Performance.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. Journal of Business Logistics, 34(2), 77-84.

Wang, N., & Katsamakas, E. (2019). A network data science approach to people analytics. Information Resources Management Journal (IRMJ), 32(2), 28-51.

Wamba, S. F., 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.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356-365.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. Journal of Marketing, 80(6), 97-121.