DEGREE PROJECT IN MECHANICAL ENGINEERING,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2020

Ideation with Big Data: A case study of a large mature firm

BHADIN BUNPUCKDEE

ÖMER TEKBAS

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF INDUSTRIAL ENGINEERING AND MANAGEMENT

Ideation with Big Data: A case study of a large mature firm

Bhadin Bunpuckdee
Ömer Tekbas

Master of Science Thesis TRITA-ITM-EX 2020:405
KTH Industrial Engineering and Management
Machine Design
SE-100 44 STOCKHOLM

KTH Industriell teknik och management

Examensarbete TRITA-ITM-EX 2020:405

Idégenerering med Big Data: En fallstudie i ett stort moget företag

Bhadin Bunpuckdee
Ömer Tekbas

Godkänt	Examinator	Handledare
2020-06-11	Sofia Ritzén	Sofia Ritzén
	Uppdragsgivare	Kontaktperson
	Anonym	Anonym

Sammanfattning

Big Data har under senaste åren fått mycket uppmärksamhet. Utvecklingen av olika teknologier har möjliggjort att en stor mängd data kan behandlas och förvaras enklare. Detta har gjort att företag har funderat över hur Big Data kan vara värdeskapande. Däremot är det inte självklart att Big Data automatiskt genererar affärsmöjligheter; företag måste förstå hur man ska förädla data och implementera insikterna. För att möjliggöra detta måste nya kompetenser införskaffas och företag måste anpassa sig till en mer medskapande arbetsstruktur.

Detta arbetets ändamål är att undersöka vilka innovationsprocesser en avdelning med data-experter som jobbar tvärfunktionellt i en organisation använder för att idégenerera för nya affärsmöjligheter. Målet är att ge rekommendationer hur företag kan bli mer effektiva vid idégenerering. Denna fallstudie utfördes för ett stort etablerat företag inom revision och inom en avdelning med expertis inom dataanalys, automation och artificiell intelligens. Datan i denna rapport införskaffades genom interna intervjuer från på avdelningen A.

Fallstudien resulterade i rekommendationer på vad som behövs att ha i åtanke vid idégenerering med Big Data. Viktiga aspekter att överväga är att Big Data möjliggör medskapande och därför är det ytterst viktigt att kunder, domänexperter och Big Data experter idégenererar tillsammans.

KTH Industrial Engineering and Management

Master of Science Thesis TRITA-ITM-EX 2020:405

Ideation with Big Data: A case study of a large mature firm

Bhadin Bunpuckdee
Ömer Tekbas

Approved	Examiner	Supervisor
2020-06-11	Sofia Ritzén	Sofia Ritzén
	Commissioner	Contact person
	Anonymous	Anonymous

Abstract

Big Data has in recent years gained much attention and interest from organizations. The rise of recent technologies has enabled data to be processed and stored in a simpler manner, thus asking organizations what value Big Data can bring to the organization. However, collecting Big Data does not automatically generate business opportunities; organizations need to understand how to process Big Data and how to implement the insights. To enable this, new competences are needed, and firms need to adapt into more co-innovated constellations.

The purpose of this study is to investigate what innovation processes a team data-expert team working cross-functional uses to ideate possible business opportunities. Furthermore, the aim is to propose recommendations of how an organization can become more efficient when ideating. The case study was carried out for a large established company within Auditing and more specifically in a support department with expertise in data analytics, automation and artificial intelligence. The data was collected through internal interviews within the department, Department A.

The case study resulted in recommendation of what to consider when ideating with Big Data. Key aspects to consider is that Big Data enables co-innovation to prosper and therefore conjoining customers, domain experts and Big Data experts is crucial for successful Ideation. Moreover, an understanding of different innovation aspects will thus help organizations understand how to ideate with Big Data more efficiently.

Acknowledgement

We would like to thank ourselves for keeping the motivation up during Covid-19. By pushing and having continuous discussion we kept the bike moving. Moreover, we would like to thank our friends and families for unconditional love and support.

Furthermore we want to give special appreciation for our supervisor at KTH Royal Institute of Technology Sofia Ritzén. Her knowledge and guidance elevated our work and helped us tremendously through the process.

This thesis would not have been possible without Levent Karaman, who gave us the opportunity to work on this thesis.

Last but not least we would like to praise our interviewees that have contributed with their experience.

TABLE OF CONTENTS

1 INTRODUCTION

In this section the background, the problematization, the purpose, limitation and the research questions will be presented.

1.1 Background

The rise of IoT and Industry 4.0 has led many companies to search for ways to understand the information from Big Data but also how to leverage those insights into business opportunities. Data is constantly produced from more utilities than ever before. Currently, the information from computers, phones, government records, health records, social media, climate sensors, and street sensors are contributing to the large upraise of Big Data. The generation of data will continue to increase as people become more dependent on technology. According to Cisco VNI report from 2015, the mobile data in the past 15 years grew 400-million-fold and Cisco anticipates that in 2020 the mobile data traffic will be 30.6 exabytes (Cisco, 2015). The growth of Big Data has contributed to the ignition of new methods of using algorithms, software, cloud and inexpensive computing power (Kiron, 2016). Traditional ways of storing data and dataset in different databases have become obsolete.

Nevertheless, Davenport (2012), states that the optimal utilization of Big Data can reduce cost and improve time-consuming computing tasks drastically (Davenport 2012). Furthermore, efficient usage of Big Data can also allow organizations to develop innovations extended from their core business by combining different data. For example, Wang (2012) found that big scale mobile phone data can contribute to a better understanding of traffic patterns for city planners and engineers. Thus, creating a better foundation for designing road networks that could decrease congestion. Moreover, Wamba et al. (2015) state that the effective utilization of Big Data will be the domain in which organizations must compete in.

The possible benefits of Big Data and a data-driven culture have not gone unnoticed for organizations. According to IDC (2018), the worldwide revenues for Big Data and Business Analytics (BDA) solutions will reach $260 billion in 2022. The momentum that Big Data has gained in recent years several recent studies have started to examine the managerial capabilities and technological infrastructure needed to implement before being able to transform Big Data into value (Sumbal et al., 2019). However, a limited amount of studies has examined the actual Ideation stage of Big Data. (Vanauer et al., 2015)

1.2 Problematization

As organizations are investing more resources into understanding how to create value through Big Data, organizations need to understand how to structure Ideation for it. Big Data is not valuable without an understanding of what the data stands for and what problems need to be solved. It requires a team of different capabilities and expertise to reform Big Data into valuable insights. Although there are several studies discussing the organizational and the technological requirements when handling Big Data, there is limited research investigating Ideation stage when using Big Data. (Vanauer., 2015)

1.2 Purpose

The purpose of this study is to investigate what innovation processes a team of data-experts who work cross-functional uses to ideate possible business opportunities. Furthermore, the aim is to propose suggestions of how to become more effective in Ideation.

1.3 Research Questions

For the purpose to be fulfilled two research questions have been formulated, namely:

RQ1: What challenges arise for the organization when ideating with Big Data?

RQ2: How can a better understanding of Ideation at an early stage in a project help organizations utilize the benefits of Big Data in a more effective way?

1.4 Commissioner

In this subchapter, the information regarding the case company and the case department will be presented.

The company, Company A, is considered one of the big four in accounting and auditing. The services range from advisory, internal audit, taxation and risk assessment. The organization conducts businesses in 15 different industries such as banking, production, retail and healthcare. Customers range from state-owned organizations, big private companies and SME (Small-Medium Enterprises). It is a global company with a large network reaching over all 6 continents.

The thesis is carried out at a support team working cross-functional with expertise in data-driven technology, Department A. The role of the department is to be able to support other domains such as advisory, taxation and audit, in creating additional value for the clients. The department was initiated from a global perspective to help the organization sustain relevance with the rapid advancement in technology. The services that Department A offers are in data analytics, intelligent automation and artificial intelligence.

1.6 Delimitations

The thesis was conducted over a period of 20 weeks. Moreover, as the field is broad it will not dive deep in the technological challenges or organizational challenges. However, it will provide insights to a foreground of how Ideation through Big Data can appear. During the 20 weeks of the process an outbreak of Covid-19 in Sweden caused limitations within data collection. More than half of the interviews were conducted through Microsoft Teams, therefore, generating fewer interactive sessions between the interviewees and us. Moreover, a field study of the organization's innovation and Ideation hub was cancelled, losing valuable observation input for the thesis.

2 THEORETICAL BACKGROUND

In this chapter the theoretical background will be presented. The definition and the different implications of Big Data will be presented. In addition, the holistic view of the innovation process from a theoretical perspective will be introduced.

2.1 Big Data – The Definition

The hype of Big Data has made the term ubiquitous. Academia, the industry and media have no conjoined definition of the term causing ambiguity (Ward & Barker, 2013). De Mauro et al. (2016) tries to formalize the definition of Big Data, in the paper "A formal definition of Big Data based on its essential features". This paper discusses four critical themes of Big Data, namely, *Information*, *Technology*, *Impact* and *Methods*.

- *Information* - De Mauro et al. (2016) explains that the upraise of Big Data is due to the massive amount of data being generated. In the early 1990s, digitization became widely popular where the initiation of Optical Character Recognition (OCR), converted traditional libraries into machine-readable files (Coyle, 2006). The transformation of analogue signals to digital ones simplifies the organization of datasets. Prescott (2013) suggests that digitization of books can be seen as "an early encounter with Big Data". Moreover, De Mauro et al. (2016) states that the availability of personal devices connected to the internet is a significant aspect too for the quick expansion of Big Data.

- *Technology* - The technological issues that arise with Big Data are the computational and storage requirements that traditional IT-systems cannot manage. HFDS (Hadoop Distributed File System) and Mapreduce are two systems developed by Google that allow the data to be split into parallel remotely located machines, hence offering unlimited distributed storage. This enables organizations to grow without considering an on-premises solution, reducing cost substantially (Tivari, 2015; Shvachko, 2010).

- *Methods* - The complexity of Big Data overrides processing it by common statistical procedures. Big Data analytics consists of a list procedure such as cluster analysis, natural language processing, machine learning (Chen et al., 2014). Consequently, leading to the new demand of expertise, a data scientist. It is a person "who understands how to fish out answers to important business questions from today's tsunami of unstructured information" (Davenport, T & Patil, D, 2012).

- *Impact* - The implication of Big Data will have a huge influence on several aspects of society. Raghupathi & Raghupati (2014) implies that Big Data has a huge potential to improve healthcare quality such as disease surveillance and decision support. Moreover, Sap (2014) gives an example of Globus, a department store chain. Globus is able to reduce 98% of the time employees spend on gathering the right information on more accurate inventory decisions with the help of Big Data and in-memory analytics.

2.1.1 The five V:s

The conjoint analysis of the four critical themes of Big Data presented above by De Mauro et al. (2016) contributes to a proposed definition of Big Data, namely, the 5 Vs. This definition of Big Data seems to be widely accepted to be a proper definition of Big Data, Volume, Variety, Velocity, Veracity and Value (Chen et al., 2014; Lau et al., 2016, Gandomi & Haider, 2015). Below follows a more detailed explanation:

- Volume
  - The size of the Big Data is tremendous. Big Data is not measured in gigabytes but rather terabytes and petabytes. The enormous amount of data makes traditional storage and analysis obsolete (Sagiroglu & Sinanc, 2013).
- Variety
  - The heterogeneous structure and sources of data contribute to a variety of data. Information could be gathered from social networks, phones, sensors, machines and much more (McAfee & Brynjolfsson, 2010). Big Data can have different structures such as unstructured, semi-structured and structured (Gandomi & Haider, 2015). User-generated content found in social media, such as comments, images and videos are common examples of unstructured data (Lau et al., 2016). Tabular data such as spreadsheets or relational databases are forms of structured data. (Gandomi & Haider, 2015). Semi-structured data does not strictly comply with strict standards but could contain tags that simplify the process of machine readability (Gandomi, A & Haider, M, 2015; Sagiroglu & Sinanc, 2013).
- Velocity
  - Phones, sensors, machines and the overall digitalization of gadgets creates Big Data at a high rate. Thus, being able to cope with the high stream, Big Data has led to an increased need of being able to analyze in real-time (Gandomi & Haider, 2015). Live personalized offers creating value for the customers from the previous purchasing pattern are cases most consumers are familiar with.
- Veracity
  - Logically, Big Data with volumes consisting of terabyte and petabyte comes with ranging quality and validity (Lau et al., 2016, Gandomi & Haider, 2015). Being able to filter and find quality data is, therefore, naturally a big concern for organizations. Utilizing different methods of data analytics after filtering out low-quality data could create more accurate business insights (Lau et al, 2016).
- Value
  - Furthermore, the large amount of raw data which is produced has a rather low value relative to its volume for a business. The data should be processed and analyzed to create understandable insights that could be translated into actionable knowledge (Lau, 2016). The different IT infrastructures need investment to be implemented thus, companies will be extra keen to see the return of those investments (Ishwarappa & Anuradha, 2015).

2.2 Big Data Analytics (BDA)

Big Data analytics describes the skills needed to analyze insights from Big Data (Gandomi & Haider 2015). Agrawal et al. (2011) remark five phases in the BDA pipeline, acquisition, extraction/cleaning, integrating/aggregating, modelling and interpretation. Moreover, Gandomi & Haider (2015), presents a more in-depth description of tools for BDA:

- *Text analytics* - is the technique of extracting information from text. The different texts openly available on the internet can with statistical analytics, computational linguistics and machine learning generate meaningful descriptions of human-generated text.

- *Audio analytics* - the process of gathering information from unstructured audio files. By using two common approaches of Transcript-based approach or phonetic based approach sound can be transformed into structured data that can be examined.

- *Video analytics* - video content analysis (VCA) is the method of capturing information from videos. There are two common methods, Server-based Architecture and Edge-based Architecture. The first has centralized servers that analyse the captured data. The latter have raw data analyzed locally on each separate camera.

- *Social media analytics* - grasping the unstructured and structured data from the different channels on social media. The user-generated content can be further subcategorized into content-based analytics and structure-based analytics. Content-based analytics could be customer feedback, product review or any data that users upload on social media. Structure-based analytics investigates the different relationships the different participants have in a different social network.

- *Predictive analytics* - the technique that predicts future happenings with the help of different statistical methods. The usage of historical data and current data is needed to be able to develop predictive analytics

2.2.1 AI, Machine Learning and Cloud Computing

In order to process and filter out Big Data, the role AI and Machine Learning have become dominant. Monotonous tasks that require routine and defined processes such as in which databases data can be carried out by computers through Machine Learning and AI. AI and Machine learning can develop deeper analysis of Big Data such as recognizing patterns and anomalies faster and more accurately. Thus, expanding the technology in innovation processes (Kakatkar et al., 2019).

Machine Learning is “programming computers to optimize a performance criterion using example data or past experience.” (Alpaydin, 2011). There are generally three types of machine learning classifications; supervised learning, unsupervised learning and reinforcement learning (Alpaydin, 2010). Supervised learning is when there are input and output variables that are assumed to be dependent (Alpaydin,2011). Unsupervised machine learning is when the aim is to take only one set of variables from an unknown probability density and estimate it. Reinforcement learning is when there is no training set, however the computer gains data by trial and error while doing a task. (Ibid)

2.2.2 Big Data – GDPR

The General Data Protection Regulation (GDPR) entered into force in May 2018 due to the upraise of digital technologies. GDPR regulates the collection, storing and processing of data, more

specifically, personal data for companies obtaining data from European Citizens. Hence, affecting every global organization, not only companies within the EU. The personal data is data such as social security number, location data, names phone numbers and ip-addresses, generally any data that can be traced back to an individual (Gruschka, 2018; Li et al., 2019). Consumers have more control and must give consent for their data to be processed (Tankard, 2016).

GDPR core principle is to embed "data protection by design and by default" into companies' culture (Goddard, 2017; ITGP Governance Privacy Team, 2016). Moreover, organizations must be transparent, providing individuals in an accessible way comprehensive information of what that data has been used for and its effect when requested. Hence, a company's IT-architecture and design must ensure that the firm's business operations always comply with GDPR (Goddard, 2017).

Furthermore, this proposes some challenges for utilization of data and therefore, Big Data as it must comply with GDPR six general principles according to ITGP Governance Privacy Team (2016), namely:

- Fairness, Lawfulness and Transparency
  - Is the usage of data where the controller is open and honest about its identity. Moreover, the controller must handle the data in a way the consumer would expect and convey clearly that the personal data that is collected will be used. (Datainspektionen, 2020)
- Purpose Limitation
  - Organization must remark what the data will be used for and scope the usage to meet the purpose. Consumers should be provided with clear information of the processing involved. However, the regulation does authorize additional processing "for achieving purposes in the public interest, scientific or historical research purposes or statistical purposes". (Ibid)
- Data minimization
  - The data that is not relevant to the purpose should not be processed. Organizations should not hold more data than required. (ibid)
- Accuracy
  - The data should be accurate and up to date. Organizations should continuously update and filter incorrect data. (Datainspektionen, 2020)
- Storage limitation
  - If the data is not needed anymore, it should be deleted. By defining a purpose for the data collection, the process of excluding irrelevant data is simple. (ibid)
- Integrity & Confidentiality
  - Personal data must be classified as confidential data. Moreover, personal data must be processed "in a manner that ensures appropriate security of the personal data, including protection against unauthorized or lawful processing and against accidental loss, destruction or damage" (Datainspektionen, 2020).

2.3 Innovation – The Definition

This chapter presents an overall explanation of innovation from a theoretical perspective. Insights on subjects within innovation is used to understand it more concretely.

Innovation is today a hot topic within the research and business world. Many divisions within companies use innovation in their names to pinpoint that innovation is within their strategies. Important to mention is that defining innovation correctly is important. The engineering, marketing and economics world all have their own spins of an innovation (Garcia & Calantone, 2002). There are several definitions of innovation, but a common definition proposed by Crossan and Apaydin (2009) is;

"Innovation is: production or adoption, assimilation, and exploitation of a value-added novelty in economic and social spheres; renewal and enlargement of products, services, and markets; development of new methods of production; and establishment of new management systems. It is both a process and an outcome."

This definition captures a wide meaning of innovation. It covers both internal aspects of innovation as processes used to create value in strategies and products and at the same time covers the exploitation part of innovation (ibid).

In order to classify innovations, since not all innovations are the same, the two that fit in this thesis have been identified as incremental and radical innovation. To stand out and be competitive in this intense market, companies need to continuously be innovative and create new advantages that creates value for the customers (Crossan & Apaydin, 2009). This can be achieved by focusing on core business but at the same time managing things in the radical area. Radical innovation is considering innovation that is new and radical whereas incremental is smaller improvements of existing products (Dewar & Dutton, 1986).

2.4 Innovation Process

2.4.1 The Front End of Innovation

The Front End of Innovation (FEI), also called the Fuzzy Front End (FFE) is the earliest stage in the innovation process. It consists of the activities that come before the formal and well-structured New Product and Process Development (NPPD) or Stage-Gate process (Koen et al., 2001). FEI is a well-known concept but there are still question marks about how to manage it. Koen et al. (2001) mentions that FEI is chaotic, unpredictable and an unstructured activity, yet an important part of innovation since it sets a common ground for further work.

As mentioned earlier, FEI is the first stage of the NPPD process. This phase consists of exploration in terms of discovery of an opportunity or an idea, before it goes to the NPPD stage where it gets developed (Eling & Cornelius, 2017). This means that FEI is all about activities just as opportunity identification, problem identification, analysis within market or technology, idea generation, evaluation and testing (Khuruna & Rosenthal, 1998).

Two important findings from a study made by Khurana & Rosenthal (1998) explains how the FEI phase could be best handled. The first is to approach FEI from a holistic view. That makes it easier to connect the product strategy, business strategy and product-specific decisions. Secondly, depending on the circumstances such as product, market, organization etc, the management of FEI can be adopted for most efficiency (Khuruna & Rosenthal, 1998).

Koen et al. (2001) emphasizes five elements within FEI that are useful for getting the scope of FEI as;

- *Opportunity identification* - This is the stage where opportunities are identified that companies want to pursue. It can be a technological or business opportunity that is identified. The opportunity is identified with the help of market and industry analysis or a lack within the company (Koen et al., 2001). The opportunity could be within the business strategy but likewise a completely new direction for the company. It is also useful to utilize design thinking and especially the emphasis stage, which will be explained further in the report.

- *Opportunity Analysis* - This stage consists of the analysis and assessment of the found opportunity. Depending on the opportunity and its attractiveness, the effort for the analysis varies. What needs to be analyzed is how well it fits the business strategy, the size of the development and the potential risks. This stage can be a part of a formal or iterative process in the companies (Koen et al., 2001).

- *Idea Genesis* - Genesis is the concretized idea that represents the evolution of the opportunity. It goes through several iterations of studies and examinations to mature it. Contact with potential customers can be conducted in this stage to understand customer needs and develop potential undiscovered needs (ibid).

- *Idea Selection* - In many businesses there are a lot of opportunities and evaluated ideas that emerge. In order to achieve the most efficient and valuable ideas, an idea selection process is needed to best handle the product portfolio (ibid).

- *Concept and Technology Development* - The fifth and final element consists of the development of a business case connected to the idea. Industry analysis, market potential, customer needs and competition are important factors that decide the size of this step before the idea proceeds to the NPPD (ibid).

2.4.2 Ideation

A company's sustaining success in a competitive market depends strongly on the first stages of innovation, where ideas are generated are crucial for successful product innovation (Koen et al., 2001). Ideation is the creative process of generating, developing and communicating ideas and finding opportunities. Ideas are the essence of society, corporation and the individual (Graham & Bachmann, 2004). Ideation helps with realizing the ideas and creating value out of them. Most ideas do not survive for long. The reasons can be lack of funds, resources and markets, but generally, the ideas die because of lack of creativity and people who thought of an idea do not know what to do with it (ibid).

In order to overcome this, involving users in the Ideation stage has been a vital contributor to product or service success (Alam, 2002; Kristensson, Gustafsson & Archer, 2004). Generally, the users can be split up to Lead users and Ordinary users. Lead users are the leading-edge users, thus persons with strong domain expertise and in-depth understanding of that specific field. Scholars

mention that lead users have a tendency of developing more radical and functional innovation compared to company innovators. (Morrison & von Hippel, 2000).

Moreover, previous literature argues that involving ordinary users in Ideation stage is not beneficial. The ordinary users lack knowledge and understanding of the technology and the surrounding the business is mentioned to hamper the Ideation. Hence, ordinary users have difficulties seeing the larger picture and the potential of the technologies (Bennett and Cooper, 1979, 1981). On the contrary to the findings from Bennett & Cooper (1979,1981), Magnusson (2009) suggest that the ordinary lack of knowledge allows them to think beyond the domain logic and therefore be more creative. However, the study does state that ordinary users do need some knowledge of the technology in order to ideate and propose suggestions that are technically feasible (Magnusson, 2009). Moreover, Magnusson (2009) findings shows that guided users, those that received relevant technological understanding of the topic produced different types of ideas compared to pioneering users (those that did not receive it). Thus, suggesting that the user's knowledge has an effect on contributing to incremental and radical new ideas.

2.4.3 Design Thinking

Design thinking is a methodology and a mindset with its product being innovation. In today's business, innovation has become a survival strategy where it is not only a new product but also new processes, services, collaboration etc. (Brown & Katz, 2011). Brown & Katz (2011) mentions that design thinking is a process of meeting technological and strategic needs through human-centered design. The emphasis stage is the first part of design thinking, where the work of understanding the need is performed. Since the problem being solved is rarely only your own, processes to understand the target are utilized. The environment becomes a crucial factor for observing and understanding what feelings and thoughts people have, and that in turn will help you to learn what is needed (Norman & Verganti, 2014).

2.4.4 Ambition Matrix

As innovation is difficult to measure, a lot of companies struggle to be innovative in practice. Therefore, the innovation Ambition Matrix was developed as a tool to make it clearer for companies about what parts of the strategy are innovative and not. This way it makes it easier for companies to illustrate the innovation goals, shown in Apendix B. A portfolio approach to innovation is how firms should allocate their resources depending on the circumstances (Nagji & Tuff, 2012). According to studies made by Nagji & Tuff (2012), firms that have a 70-20-10 allocation of the ambition strategy outperforms the competition. Roughly 70% should be invested in core business, where investments are made on safe bets where the company is within their comfort zone. Furthermore, 20% should be allocated on adjacent businesses that are less safe. The final 10% is allocated to transformational and disruptive innovation that has higher risks.

Nagji & Tuff (2012) mentions in their study that;

"*Rather than hoping that their future will emerge from a collection of ad hoc efforts, smart firms manage for 'total innovation'*".

To manage the full control of total innovation, companies need to find a common language to describe innovation from top management and down. Leaders need to communicate that innovation is something that the company has within their strategy and to manage to be competitive, they have

to do things differently, together with different people, different motivational factors and innovation processes (Nagji & Tuff, 2012).

2.4.5 Diversity

Sourcing of knowledge through all possible ways is crucial for achieving innovation. There is a growing interest in research within open innovation and how diversity within a firm can create value in terms of innovation. One important way is to have a diverse culture within the company and especially within teams (Bogers et al., 2018). Bogers et.,al (2018) discusses how diverse educational backgrounds and varied work experience in companies can affect how open a firm is to external knowledge. Research shows that firms use different ways for sourcing external knowledge, such as suppliers, customers, universities, collaboration and by employing people with different backgrounds. Dahlander & Gann (2010) mentions that the starting point of being innovative as a single organization is to avoid isolation. Companies must engage with other types of partners to get new ideas. This is crucial to stay competitive in the market.

As mentioned above, diversity within teams is important for staying competitive. Two different types of diversity will be mentioned in this section as work experience and educational background. People with different backgrounds can easily get in touch with their network to potentially create value and most importantly innovation (Ahuja, 1998). Building an external network for future collaborations is highly connected to previous connections. People with different types of educational backgrounds and work history can easily support each other concerning different problems.

Bogers (2018) discusses that there is a positive interaction between the size of external knowledge that could broaden the company's knowledge base and the educational and work experience diversity. Employees with diverse work history can easily connect functional backgrounds due to their experience and capacity. Since they usually have worked with a lot of projects, mistakes can be avoided, and effective workflows can be implemented. With people with diverse educational backgrounds, companies are getting use of integration of diverse knowledge and perspective in topics that can complement and support other teammates (ibid).

2.4.6 Management Issues

Dodge et al. (2017) mentions that culture can affect performance between 20 and 50 percent on organizations within the same industry in a positive way. Leadership behaviors are crucial within culture, thus this can hinder innovation if not correct behaviors are established.

Management within a firm may have different opinions towards innovation and how to set up the strategy with the help of innovative mindsets (Deward & Dutton, 1986). They can be conservative and against changes or encourage them. However, if the power is in their hands, upper management converts the attitudes within a firm to action plans.

A study made by Dodge et al. (2017) shows three important findings on what employees perceived as an innovative organization;

- Providing organizational encouragement
- Ensuring challenged work

- Fostering support within the work group

However, the importance of these dimensions depend on how the companies are organized. Three different organizational styles mentioned in the study are high *level of control by management and high/low level of support to employees*. In organizations that have a high *level of control,* it is more important that management is *ensuring challenged work.* Companies that have a *high/low level of support to employees* need to provide *organizational encouragement* in order to have the greatest positive impact on innovation. However, it is important to *ensure challenged work* to employees that have a *low level of support* since this increases innovation (Dodge et al., 2017).

2.4.7 Innovation as an Outcome vs as a Process

The difference between innovation as a process and an outcome is often blurred (Crossan & Apaydin, 2009). Innovation is defined as both an outcome and a process. Innovation is the first phase of a new product, process, method or system (Quintane et al., 2011). Innovation as a process is described by Quintane et al. (2011) as "bringing into use" or "development and implementation" of an idea. There are two phases for the innovation processes; idea generation phase and the implementation phase. The idea generation phase includes the creation of something new, often initiated in new business or market opportunities to be able to be competitive. The implementation phase is where the experiments are done and mistakes are found (Quintane et al., 2011). This way, ideas are tested for integration and viability for a specific purpose (Ries, 2011). Both phases are connected since an increased amount of ideas also means potential increase of implementation.

Innovation as an outcome is described as a product, process, software, idea etc., that is considered as something new for the environment it is established in. The novelty of the outcome is graded on the degree of innovation. Crossan & Apaydin (2009) discusses as mentioned earlier, an innovation as incremental or radical. Not all innovations are novel. Many organizations have a misunderstanding of innovation. A technology that was an innovation when it was introduced years ago is not necessarily considered as innovation today. Firms offer innovative services and products that become the new normal in the future. Therefore, to have the mindset of innovation as a process is not always the key to success. To be able to have long term competitiveness and stand out with the offered products, the whole company needs to have innovative outcomes that enhance the business strategy (Crossan & Apaydin, 2009). In short, innovation as an outcome is the degree of novelty of a product or process (Quintane et al., 2011).

2.5 Big Data and Innovation

Scholars have in previous studies been clear that Big Data can develop innovation in several different fields, from process to products. A data rich environment is said to be critical to an organization's innovation activities (McAfee & Brynjolfsson, 2012). Rindfleisch et al. (2017) proposes two separate ways of innovation using data, that is, Innovation from Data (IFD) and Innovation as Data (IAD). Innovation from Data is the usage of insights to provide solutions to customers (Bharadwa & Noble, 2015). Moreover, in IFD customers are rather passive. Instead, organizations take a more active role by example observing consumer online behavior to produce more developed products for the customer (Rindfleisch et al., 2017). On the contrary to IFD, IAD invites consumers to be more proactive. The consumers are the most relevant actor and the firms playing a less relevant part in the innovation. Instead of acquiring, analyzing and developing on the data, organizations could just be the data provider (ibid).

Moreover, Lozada et al. (2019) proposes that Big Data can influence the co-innovation process positively. Co-innovation is the participation of different stakeholders in the company including external parties for creating new innovation. Lee et al. (2012) states that "if an organization is interested in only pursuing its own goals without any regard to the shared value with its stakeholders, it will not survive long". The tremendous amount of Big Data enables resourceful inputs identifying possible insights such as product value or profilication. Thus, decisions can be based on data that can improve the possibilities of product acceptance in the market (Zhan et al., 2017). Furthermore, BDA enhances the collaboration between different stakeholders in co-innovation in regard to solution, problem and challenge definition. Hence, Big Data can provide insightful information to the co-creation (Lozda et al., 2019).

2.6 Big Data and Ideation

There exists very limited scientific research regarding the Ideation stage of utilizing Big Data (Vanauer et al., 2015). Several case studies from consultancies firms have been published discussing the impact of Big Data (Court,2015; IBM, 2012). Moreover, Vanauer et al., (2015) conclude that most research present generic methods of Ideation, assessment and implementation of Big Data. Moreover Vanauer et al., presents (2015) two different views of Ideation with Big Data, the business first approach and the data-first approach.

Business First (BF) approach is initiated by understanding the business requirements and assuming that those requirements could be tackled by the help of Big Data. Vanauer et al., (2016) discusses that the main field of where BF is beneficial is where improvement can be done on existing processes. The first step is suggested to be the identification of business requirements, organization goals and known operational challenges (ibid). The paper argues that for BF one of the significant prerequisites is the domain knowledge of the operation process and concludes that first/mid-level management and personnel are the key innovators. Additionally, by incorporating data scientists and data analysts a discussion regarding the development of operations could be commenced.

On the contrary to BF, Data First (DF) is focused on advancing services to be sold to other entities by examining the data. DF can be seen as more revolutionary as it is suggested to focus on producing a new business model. One of the primary tools used for this is the Business Model Canvas. The Ideation approach relies heavily on the value proposition derived from the available data. Vanauer et al., (2015) suggests that defining a value proposition needs the capability of thinking "imagine what does not exist". Moreover, the authors explain that innovators in this stage must understand that the value proposition is not based on isolated data but combinations of multiple data too. Under these circumstances data might be limited in terms of availability in-house, thus finding data sources that could be bought or obtained in other ways should be proposed. (ibid)

Kronsbein and Mueller (2019) presents a Data Innovation Board used for Ideation, see Appendix D. The authors suggest that the Data Innovation board is useful as one of the earliest stages of Ideation where the participants have low knowledge of data or Big Data.

The findings from Kronsbein and Mueller (2019) depicts a simple way of concatenating data into Ideation, see appendix A. The Data Innovation Board addresses four data standpoints, namely, existing data, additional data, used existing data and data ideas. Existing data reflects upon the data that is available. Additional data is the data that is available but needs to be retrieved. Used existing data is the data that has been used already. Lastly, a data idea is the goal of which the data should accomplish for the specific need.

3 METHODOLOGY

This chapter will describe the methodology of how relevant empirical data was collected and analyzed.

3.1 Research Approach

This research paper will follow a case study approach. Eisenhardt (1989) states that case studies are well adapted for the early stages of a research topic. Big Data is still a novel theme, where its implications are not fully explored yet, especially the Ideation phase. (Vanauer et al., 2015). Beforehand, the problem definition was not clearly defined. Thus, in order to collect data to explore the possible problem within the organization and Department A, a qualitative iterative explorative research approach was conducted. The qualitative research approach investigates "the social setting from insiders' perspective and generates description analyses of contexts "(Lapan et al., 2012). Thus, enabling a suitable method of understanding existing processes, identifying problems within the Organization and furthermore a suitable proposition.

The process initiated with broad research within Big Data. This was to generate background in a somewhat new topic that could generate the first insights into suitable topics and questions. Thereafter, the first phase of interviews was with open-ended questions in order to obtain an understanding of the organizational context, process within the organization and how Department A utilized Big Data. Having open-ended questions enabled insights and new themes as an output from the interview. Generating a flexible structure that allows new fields to be discovered, common for qualitative methods (Devers & Frankel, 2000).

Furthermore, the information gathered from the first phase of the interview was transcribed and coded into themes, generating into a deeper and comprehensive literature study of Big Data relevant to Department A. Consequently, another round of semi-structured interviews was conducted, phase 2, which had a more focus on the themes gathered from the findings in phase 1 and from the literature review. The interview guides can be seen in Appendix D.

3.2 Data Collection

The case company investigated is one of the big four in Auditing. However, in recent years the organization has understood that the rapid technological advancement has opened additional business opportunities extended from their core capabilities. The first two interviews were conducted with the innovation lead and the responsible manager for data analytics to understand the overall goal and the innovation ambition context of the department. Moreover, the first two interviews allowed a better understanding of additional stakeholders to interview and reach out to. Prior research was done by searching for a broad source of data, such as internal and external documents (presentation, articles, customer examples, previous projects) published by the Department A in Sweden and abroad. Consequently, leading to a rich understanding of the context. In appendix A, a detailed explanation of data collection sources and its intended use when analyzing. Moreover, to ensure quality, interviewees were chosen by using purposive sampling, informants were asked other candidates relevant to the topic (Tongcao & Dolares, 2007). Thus, contributing to informants with rich knowledge were targeted as participants (Miles & Huberman, 1994).

3.3 Data Analysis

The interviews were transcribed and analyzed following a method of open coding. Hence, no pre-set of codes was established beforehand but rather developed and altered continuously (Strauss & Corbin, 1998). Moreover, a theoretical thematic analysis was conducted, where "we coded each segment of data that was relevant to or captured something interesting about our research question." (Maguire & Delahunt, 2017). The relevant themes were thereafter exported to external documents and summarized into key insights. Thereafter the insights of the different interviews were compared to each other following Eisenhardt (1989) cross-interview patterns. Allowing data from each interview to be analyzed separately before generalizing (ibid). The data collection and data analysis were done simultaneously allowing further themes to emerge along the interview process.

3.4 Evaluation of Methods

Qualitative research approaches are suggested to generate an understanding of the dilemma due to the time and the relationship spent with the subjects. Duffy (1986) states that the strength of a qualitative approach lies in the holistic focus, allowing a deeper and more valid understanding compared to a more rigid one. Thus, bursting quality to the data. On the contrary, quantitative is based more on numbers developing a more structured system (Carr, 1994). However, quantitative research methods have been argued to neglect insignificant data, therefore invalidity and distorted data (ibid). Following the strength of qualitative research approach scholars have suggested that it also presents a weakness of the subjectivity as researchers could have trouble distinguishing from their experience and those of their subjects (Sandelowski, 1986). Furthermore, to examine the quality of the study analysis according to Kirk & Miller (1986) two measures of reliability and validity have been investigated.

Reliability and Validity

The assessment of the reliability in this study can be done in different ways. Before the first interviews, research was done within the investigated field to get a better understanding of the problem. The interview process was divided into two phases where the first phase was to get a broad understanding of what Department A does and how they do it. Thus, open-ended interviews were held with the innovation lead and the manager responsible for the Department A to get a context of how work was done within the Department And which employees could give most insights within the investigated subject for further interviews. All interviews within the first phase were conducted with an open-ended approach in a short period, allowing consistency of the types of questions asked and also the environment in which the interviews were held. Interviews were held with 50 percent of the people in Department A, and with another department that worked together with Department A to get their perspective of the collaboration with Department A. This gave enough insights to begin preparing the second phase.

The second phase of the interview process consisted of structured interviews with questions focusing on the problems arising in the first phase that was connected to the studied field. The questions were well prepared and asked in such ways to get answers that were motivated and well explained rather than short answers. Follow-up questions were asked to dig deeper into the problems if needed. This way, answers from the interviewees were obtained and they were necessarily not aware of what factors were in mind when asking the specific questions.

Identical observations were obtained from the interviews and used for developing the results and solutions. In addition, all interviews were recorded and transcribed, and documented in an organized way to be able to get back to it easily for listening and reading for eventual validations.

Achieving validity in qualitative research was obtained through construct validity. It was addressed by using several levels of sources, such as diverse interviewees within the Department A and external people working together with the Department A. This way, partiality was avoided.

4 RESULT AND ANALYSIS

This chapter presents the findings of the study and how Department A project processes look like. Following by how Big Data and innovation helps them deliver their products in a competitive way.

4.1 Organizational Context

The organization which this thesis was conducted at is coined one of the big four in audit. The core services lie in tax, advisory and audit. Due to the rapid advancement in technology and the harsh competition a global initiative within the company has been carried out to tackle this, Department A. This department consists of expert knowledge within data analytics, AI and automation, quite far from the organization core competencies. The aim of the department is to spark innovation and digitalization for each domain in its deliverable to external customers but also for internal processes.

"Generally speaking, at some point a company will become a software company regardless." - **Interviewee 1**

Although the organization is seen as a global company, each local office is run independently. Thus, the organization's operations differ from country to country. Thus, cross-collaboration between geographical is possible in theory but rather complex in practice. However, due to Department A being a global initiative cross-collaboration and knowledge sharing is easier.

Department A can be split up into two departments, a software developer side and a business unit working with partnership with the domains. The developer side of the department works mainly with developing solutions and maintaining the software up to date. The business unit is seen more as an advisory unit helping other domains within the scope of AI, automation and data analytics. Moreover, the different projects could be categorized into two fields, internal projects and intra-external projects. Internal projects are mentioned to be projects with internal end-customers. Whilst, intra-external are internal projects but with external end-customers.

4.2 Project Process

During the interviews, the interviewees were asked to describe their work processes and in what way they received their projects. The point of views presented represents employees in Department A working with different projects.

As stated previously, Department A is mainly seen as a support function for the core domains which includes tax, advisory and audit. These domains have the client facing roles and have work experience from a variety of projects. Since Big Data is a growing topic, a lot of projects need data experts to be able to offer help to clients. As an interviewee mentions;

"Other departments come to us for support in their projects"

- Interviewee 2

Majority of the projects come from these domains which are the ones owning the relationship with the clients. They turn to Department A when facing complex technical problems where they have a lack of knowledge, specifically within Big Data management and value creation through data. The project initiation in other hand, comes from the upper management. Different leaders and partners have a broad contact in many lines of businesses. They are the ones distributing the projects to the domain experts.

"The domains get projects through our upper management."

- Interviewee 3

and

"Partners and upper management own the first relations with potential customers."

- Interviewee 4

Upper management are initiating the projects together with the potential clients. Then, the domains are usually the ones conducting the first meetings with the client where a more specific discussion around factors frustrating the clients are brought up. People experienced in defining problems and presenting ideas attend those meetings.

"Other departments own the relationship with the customers, not Department A."

- Interviewee 2

4.3 Ideation

In a later stage, when support is needed, Department A can ideate with other stakeholders. There are several ways of how Department A is receiving their projects from these domains. There are not specific processes used when letting Department A into the projects. The employees that have an understanding what Department A is capable of, are the ones Department A usually collaborates with. One interviewee mentions;

"The first initiation is us speaking internally to the people that understand what we do. The second part is them saying who this could be relevant for in their client list. And the third stage is what problems could be relevant for these clients."

- Interviewee 5

Since Department A is still a relatively new department, other domains within the firm have had hard times understanding the way they can use them as support and what their real strengths are, but this is getting better each day. Department A takes different initiations to let the whole firm understand and learn that a new support team is existing and growing within the company. An interviewee says;

"Other departments know our potential more today since we have held workshops with them, and we are trying to have continuous communication to present how we can give support."

- Interviewee 4

Department A is continuously trying to improve their communication with other departments to show leverage with their support. They conduct workshops to be able to show demos that can help them visualize products.

After getting projects from the domain experts, Department A gets a briefing on what the problem is. This stage is very dependent on the project and problem. An employee says;

"The other departments have the problem already defined, but not always."

- Interviewee 4

Many times, the problem is well defined when Department A is getting to work with the projects, but not all the time. Cases where the problem is not precise, Department A conducts regular meetings or prepares workshops with the domain experts in their innovation hub depending on the complexity of the project.

"We conduct workshops when the problem is complex enough."

- Interviewee 6

Workshops helps Department A to dig into the problem and help the domain to define a more precise problem with the help of different tools. In some cases, the client can be part of the workshop to make it easier for Department A to get answers on questions related to the problem that can be difficult for the domain experts to answer.

"We use workshops for problem definition and idea generation with design thinking as a method."

-Interviewee 4

As mentioned, a method used in the workshop is design thinking. Department A are the experts within data and digitalization; hence they need a good understanding of the problem to be able to create a solution. Design thinking helps to emphasize with the domain experts and the clients, however, there are no set processes of when to conduct workshops.

"Our innovation hub is using design thinking in their workshops together with other departments and clients when necessary."

- Interviewee 6

Workshops are not conducted for every project for idea generation or problem definition as mentioned by an interviewee, but only when it is necessary.

As mentioned earlier, Department A is not owning the relationship with the clients. They come in later in the project, which means that the majority of the cases are defined from the domain experts. Even if Department A are the experts within data and the ones that develop a solution for the clients, they dont attend the first meetings. As an interviewee says;

"I think Department A currently is not even in the second conversation, more in the fifth and tenth conversation."

- Interviewee 5

"Sometimes we just end up discussing with the internal domains, and not the end-customers." **- Interviewee 5**

This shows a lack of communication between the clients and the data experts. It was mentioned by several interviewees that they were generating new ideas and opportunities for the clients when they got to meet them in person and especially in workshops. To be able to create business opportunities and understand the clients, two interviewees mentions;

"Everyone in Department A should be involved in the first meetings with the client, in reality."

- Interviewee 5

and

"We want to be involved from the first meeting with the clients to the last. A cloud architect could attend the firsts meetings for overall comprehension and data engineers in the middle of the process for asking their technical questions and analysis."

- Interviewee 4

Employees see a clear need of Department A being involved more directly in these client meetings to develop ideas and opportunities. Several respondents mentioned that being able to have interactive discussions with clients have made it possible to open new opportunities that have not been on the agenda. However, the domains do still not have any processes within their work to let Department A in the first meetings as an interviewee mentions it;

"The other departments don't have a process to let Department A join the first client meetings"

- Interviewee 2

4.4 Innovation

Innovation is a widely used term within companies, but everyone has their spin on what it means. One employee mention;

"Innovation for us is to create an ability that has not existed for the customers."

- Interviewee 6

What is mentioned is to create a solution that is new for the customer. This does not necessarily mean that the solution is new or created from scratch as a new service. However, the respondents have different understanding of innovation.

Another employee says;

"Innovation is about releasing the time horizon from one or two years and instead look further into the future."

- Interviewee 7

The time horizon of innovation is mentioned as catching up with the technology and ensuring innovation even in the long term. The interviewee focused on using the latest technology and tools

as innovation. This to keep staying competitive and relevant in the market for their customers. The interviewee also says;

"Every company should continuously work with their business models and product portfolio to ensure relevance in the long-term competition."

- Interviewee 7

According to the interviewee, employees mix up innovation with the small incremental changes they do in their services. As classic research and development is about doing things more efficient and cheaper, people think that innovation is the same thing.

Furthermore, another interviewee says;

"Innovation for me is trying to deliver something new. It could be software or deliver perspective."

- Interviewee 8

The interviewee mentions innovation as an outcome or a process. Focusing on trying new ways of doing things that is usually not done or focusing on different outcomes with a delivered service. Offering services as automation or predictive analytics is not innovation according to the interviewee, continuing mentioning that people usually tend to fail the values and meaning of innovation.

One interviewee also mentions;

"Radical innovation for myself is that for every event we do, we do something different."

- Interviewee 5

The interviewee focuses on doing things differently, with different outcomes. Mentioning that services that are innovative today have a due date. Every idea generation for customers should be customized for maximizing the novelty in a solution according to the respondent.

The definition of innovation for Department A was not set as something clear. Many respondents had their own spin on the definition and several of the employees landed on design thinking as a method that was used to be innovative.

4.5 Big Data – The Definition

The interviewee's definition of Big Data resonates closely to each other. One of the foremost characteristics of Big Data was the volume of information. Petabytes and gigabytes of data were seemingly ubiquitous from customers both also within the organization. The endless production of data brought the discussion of how to leverage it. With the vast amount of data generated every day has shifted the technology push of the prior IT-systems.

"If I have computer on-premises and I receive a data set from a customer with 20 000 rows. No problems, I can do it on my computer. But if I get unstructured data such as 40h video or 60h audio files it will not work."

- Interviewee 10

The solution to handling Big Data can be managed by two different approaches. One is having computational power on promises. Thus, servers and processors are located within the company in a server hall where employees can access it. This was discussed as an outdated solution where scalability and cost were rather limited. The second is cloud computing, a service that allows users to have unlimited access to computing power, thus making it scalable, cheaper and faster.

"I can use cloud computing and set it up in one minute. Suddenly, I can have 6 servers which I used to compute something in 5 minutes instead of six hours."

- Interviewee 10

and

"Cloud has the same service in any part of the world, the services are scalable. It does not require much to make it right and scalable. The con of having things on-premises is that you cannot scale it wildly to another one. You must contact their IT-Department And ask for space in their data environment and look at the risk of installing it in their datacenter. It becomes a big process to take another on-premises solution to another country even within the organization." **- Interviewee 4**

However, Big Data or data, in general, is not valuable in its raw form. Understanding the intended usage or the scope of how the data should be adapted to fill its purpose is vital. Allowing Big Data to be obtained and stored within the realm of an organization does not automatically provide benefits. The mindset of enabling data to become valuable is key. The low value relative to its volume can impose difficulties for businesses as the data must be processed to develop informative insights.

"People say that Big Data is the new gold, but data is like letters if you do not make a sentence it is just letters. If you do not know what to do with the data, it does not matter".

- Interviewee 2

Furthermore, having faulty data or missing data is a key aspect that must be managed. To allow data to be massaged and adapted in order to enrich the data is crucial. The level maturity of actual quality should be continuously assessed, where the data from different sources had to be transformed from raw data in the data lake to golden data.

"It could be raw data, which consists of different levels of information. There could be faulty data and you need to wash it and massage it to be able to adapt to something."

- Interviewee 9

4.6 Big Data and Ideation

The Ideation stage initiates through finding the business requirement first. Additionally, the different cases varied tremendously on the domain that they were delivered with. Furthermore, one of the most prominent actions is to empathize when ideating.

"As an expert, it is important to develop empathy for the domain expert, so we want to shadow-follow them. We cannot transform without understanding the domain."

- Interviewee 9

However, ideating directly from the data seemed to be seen as a rather destructive approach for ideating.

"Usually it is very bad practice to start from the data because you will end up very detailed without knowing what they are looking for" - ***Interviewee 1***

Furthermore, collecting and scanning as much information is mentioned to be an important initiation of Ideation. To be able to collect data not only from the customer but identifying other relevant data sources was an aspect brought up by the interviewees. Moreover, gathering more information that could complement the data to give it some context was vital too. The multitude of data gathered was important to identify what aspect to focus on but also what Department A could use to leverage. Having access to open data sources such as data from SCB or SMHI could give a broader and more in-depth analysis, which would not be possible by prior software such as analyzing data in excel.

"The idea is that in the first stage you scan information and collect as much information from the client-side and from internal organization. What does the client know? What more information could be used to understand them better?"

- Interviewee 5

Having an exploration stage within the analysis of the Big Data was argued to be mandatory by the participants. Data should be broken up and experimented into visual diagrams such as graphs. But no framework was mentioned as to how the data should be experimented with.

"No framework is used. But you have to get a feeling of data. You look at it and it generates questions that we can ask the data-owner…you must have a data exploration part but how you look at it depends on the data."

- Interviewee 2

To be able to allow Ideation and co-innovation to develop the organization has invested in an Innovation Centre. The goal of the Innovation Centre is to gather different stakeholders and end-customers for a session for inspiring and creating a foundation of new business projects with Big Data. The Centre consists of a fuse of digital tools and manual tools such personalized software that allow the big screen to be used as a digital whiteboard.

4.7 Big Data and Innovation

The sense of how Big Data could be used to innovate and create value for the organization was closely discussed by delivering with the different domains. Having Department A as a unit delivering independently with similar clients facing roles like advisory, tax and audit was expressed as a bad business practice from the interviewees. It was mentioned that in terms of AI, data analytics and automation there are a handful of other firms with more experience. Thus, leveraging Department A competencies in Big Data would be more successful if it was co-created with the other domains.

“If we think of data analytics services, there is a tone of it-companies and data analytics companies.” ***- Interviewee 2***

By joining forces, it was stated by the participants that services could be more adapted and better offered to the customer. To be able to utilize the vast capabilities of the whole organization more solutions could be targeted, thus creating a unique offer that few organizations had the capabilities to match.

“For us it becomes a blue ocean strategy if we can say to the customer, we are experts in audit, but we can build a data warehouse explicitly made for data and your purpose” ***– Interviewee 6***

and

“By doing this we can build smart services based on our domain expert that can generate insights that the customer would not have obtained otherwise” ***- Interviewee 6***

Moreover, it was discussed that Big Data enables one to think beyond one's own capacity. Allowing, more disrupting innovation to take part. Thus, reaching a wider path of approaching a problem.

“Let's say that we have a big retail company, and with our expertise we can start to analyse all their stores live. We can go analyze each store geographically and see how they perform daily. This was not really possible before.” ***- Interviewee 6***

4.8 Communicating Big Data Opportunities

Discussing the possibilities of Big Data with stakeholders with low knowledge within was stated to be problematic by the informants. The understanding of the capabilities that could be used to help the end-customer with their businesses was sometimes undefined by the customer and internal expert domains.

“Generally, people who are not in the field of Big Data think everything is about the data but is not.” **– *Interviewee 1***

and

“*Big Data and the competence around the technical part of that is hard to discuss with someone who is not maybe that technical.”* ***-Interviewee 8***

It stated that it was not unusual the domains would reach out to Department A asking:

“We got a lot of data from the client, what can we do with it?” ***- Interviewee 1***

In some cases, the solution the domains discuss with the end customer is not always possible. Thus, allowing expectation occasionally to be higher.

“As soon as we start talking about Big Data and data analytics, Department A should be involved from the start and scope it. So, you can manage expectations but also have a conversation of something that is actually possible, sometimes impossible things are discussed. “ ***- Interviewee 5***

The interviewees remarked that it was understandable that those that did not have the domain expert within the field of Big Data had problems defining approaches. In order to mitigate obstacles, the informants highlighted the importance of allowing the requester to take a step back to understand the business problem and thereafter guide them into solutions that might exist.

Moreover, speaking theoretically around Big Data and BDA was mentioned to be rather a destructive approach when communicating. Instead, focusing more on the business impact was seen as a more constructive approach to introduce BDA. This is mainly due to the organization having clients in c-level positions. Examples were given by the different processes in accounting, where a specific file would be manually converted into a spreadsheet, where employees would cleanse, create macros and visualize this every year for each of their end-customers. Under these circumstances the technology within Big Data can automate much of the process by handling the data from different customers automatically rather than manually. Hence, allowing the repetitive task to be less manual labor. However, as Big Data enables one to think beyond regular capabilities causes the discussion to be harder to estimate and grasp.

"But many of the times to even discuss that and come to the conclusion is okay, if you do this for me, I will invest in you. You will not come to that conclusion because they (the customer) will not have a proper estimation of how much time they actually put on it… We can talk and talk, but building things takes time and the customer does not know how much they will save if you put this process for them." – ***Interviewee 8***

and

"C-level people will not care that much about the technology, because often they will not understand that. It is more about the business aspect and the technical details do not matter that much." ***- Interviewee 8***

4.9 Big Data Analytics and Technology

Big Data is stated to also refer to the technology used to handle the data. AI, Machine Learning and Automation are terms used in the same context when discussing Big Data Analytics. Azure, a platform developed by Microsoft, was considered the foundation enabling the development of Big Data analytics.

"Where we develop everything has a lot of interlinked services or microservices. You just need to get the data into the Azure cloud and then you have access to the services." ***-Interviewee 10***

However, having time to set-up a process and an infrastructure where all components are clearly structured is harder. In order to enable the first step, Department A has developed an enabler of Big Data platforms. The aim is to be able to store different data sources and include the different microservices to compute in the platform.

"The software is not a solution but a storage and computing enabler. A technology stack." ***- Interviewee 9***

Moreover, the informants stated that for successful BDA it is important the data is available for the employees to discover and explore.

"For me, access to the data within the organization is crucial for new business opportunities. I have an example from my previous work where the access to the data and the understanding of

the process opened up a possibility that saved millions for the company. If the access was not there, such an amount of money would not have been acquired by the company.” - ***Interviewee 8***

However, the organization did not have an enterprise hub where information could be accessed by the employees.

“It is a project I initiated quite recently… We do not have an open-source discovery platform which I initiated internally. - ***Interviewee 9***

5 DISCUSSION

5.1 Innovation – The Definition

As Crossan and Apaydin (2009) mentions, innovation needs to be a value-adding novelty in economic or social spheres. An ability that the customers do not have is value adding to the customers, but not necessarily to your firm. Focusing on developing processes and products internally with innovation in mind can keep the innovation work alive within the firm. Measuring innovation is difficult but setting goals and using tools for making different innovation degrees clearer can help the firm in their strategy.

5.2 The Front End of Innovation

Quintane et al. (2011) describes that development and implementation processes can be seen as factors that can be innovated. Delivering novelty is a crucial aspect of innovation and what needs to be understood as a firm. The first activities in the innovation processes are FEI. This phase consists of exploration in terms of discovering the problem or an opportunity. FEI is divided into five elements where each of them is connected to each other. The first element is opportunity identification. This is very crucial for further advance (Khuruna & Rosenthal, 1998). Conducting different meetings with brainstorming, design thinking and similar tools for evolving opportunities together with the clients is number one priority. If data-literate people also attend these meeting since they are the solution developers, more opportunities can come about. What happens at Department A is the opposite way of creating customized meetings and opportunity identifications. When the problem is pre-defined from the domain experts, opportunities to create something outside the box together with people who develops the solutions gets limited.

Department A as a support team, gets in the projects after the fifth element of FEI, concept and technology development. This stage consists of developing a business case connected to an idea. Several stages where customer needs are satisfied are already completed, which means that the FEI process is not a stage where Department A is highly involved in.

5.3 Design Thinking

The first stage of idea generation and problem definition is as mentioned, very important. Understanding the customer and emphasizing with them using methods like design thinking is crucial to innovate and meet customer needs. Brown & Katz (2011) is pushing on emphasizing with the customer to be able to understand them. Department A is not being involved in employee meetings where the problem identification is made. This causes problems in later stages, where employees realize that being involved in meetings with the customers is crucial for defining precise problems and being able to generate new ideas. This is also very crucial for innovation. Not being able to ask questions directly to the customers can cause misunderstandings that can have further impacts on the projects. Innovation as a term is included in Department A:s strategy as; *Innovate with, deliver with and sell with.* This includes that it is purely a support department. As a support team, it can be difficult to understand the problems the clients' faces if Department A is not involved during the first meetings with the client. Design thinking as a method is a good way to start

discussions and brainstorming together with the clients. Important is that Department A, as a support team, can be a part of the design thinking phase together with the end clients.

5.4 Diversity

Bogers et.,al (2018) discusses how diverse educational backgrounds and varied work experience in companies can affect how well they perform in terms of new knowledge and innovation. That, in combination with engaging with different types of partners and not being isolated in your own workflow opens up lots of innovation opportunities. Ahuja (1998) mentions that in order to support each other efficiently, teams should be diverse in terms of educational background and work experience. When domain experts are having meetings with the clients for defining the problem, it is crucial that people from Department A are pushing to be in those meetings. The first meetings are where the opportunity identification stage is held. In order to be creative and maximize the ideas, data-literate people, which are developing the solutions, should be prioritized to meetings to support the others. Domain experts developing the projects and getting help from Department A are often people with the same background and knowledge. It can be difficult for those people to know what potential the support team and data-literate people have when developing the solutions. Therefore, diversity within the team when meeting clients should be prioritized for best understanding of the situation and minimize misunderstandings. Many employees mention problems that are rising because of the fact that they are not participating when identifying business opportunities. Pre-defined problems are changing when Department A influences the domain experts and new meetings are held with clients. This stage is completely unnecessary and can be solved by letting all people involved in the projects be a part of every important stage including identification of needs and problems.

5.5 Ambition Matrix

The Ambition Matrix that Nagji & Tuff (2012) shows in their study can be used for further understanding of the type of innovation. To have long term goals with radical innovations and short-term goals with incremental innovations can keep the strategy up to date. Using the latest technology to deliver services was important for the company to stay competitive. That way, they managed to do incremental changes to their services that made a difference in the outcome. This shows that setting innovation goals give results. Department A had no innovation goals set from upper management. The only goal was to use the latest technology for their services. Innovation is more than just services as mentioned earlier. Department A can set goals within their processes of working and strategies used to identify needs and solutions. Integration of an innovation portfolio requires the right skill. In most firms, the majority of employees working with innovation are stuck to projects that are core offerings. This hinders the employees to be creative and think out of the box to deliver new and radical ideas. Samsung shares a finding where they have studied that people involved in radical and transformational innovations are most efficient when they are separated from the core business, either organizationally or financially. Without that distance, they will not be able to be at their peak creativity because of influence (Nagji & Tuff, 2012). However, for many companies, innovation remains as activities that miss its own processes. It is not seen as a source of energy where it could be implemented in the overall strategy to drive a growth.

5.6 Management Issues

In order to implement changes, upper management needs to understand why it is important to be able to get through changes. Dodge et al. (2017) mentions that culture can affect organization performance between 20 and 50 percent. This means that upper management are very important in decision making in firms. Leaders need to encourage employees and listen to them. One important finding is that fostering support within the work groups is perceived as very important for innovation by employees (Dodge et al., 2017). If Department A has problems attending the first meetings with the domain experts, upper management should support them in order to try the changes the employees propose. Examples given by interviewees includes opportunity identifications that have led to innovative projects and products because of ensured support by upper management.

5.7 Big Data – The Definition

The definition of Big Data stated by the informants coincided with the findings from the literature. The essential features presented by De Mauro et al. (2016) were continuously mentioned as aspects that describe Big Data, one of the most prominent features was the technology surrounding Big Data. The interviewees mentioned clearly that the technologies such as the Azure platform has allowed Big Data to be utilized in many promising ways. Moreover, the new technologies have, however, positioned others not within the field to have a low understanding of what can be done.

Furthermore, the five V's proposed as the definition by (Chen et al., 2014; Lau et al., 2016; Gandomi & Haider,2015) of Big Data are continuously mentioned by the informants when trying to describe Big Data. The enormous amount of data made continuously everyday with different structure and validity proposed big challenges for organization.

5.8 Big Data and Ideation

The empirical findings from the research supports the theory regarding BF being a common Ideation concept when improving existing processes (Vanauer et al, 2015). It is clearly discussed by the informants that understanding the business requirements is the foremost important factor. Similarly, to what is suggested by Vanaue et al. (2015) Department A identifies objectives and tasks to be improved. Thereafter, senior managers and/or domain experts set up objectives that departments A, BDA employees must meet. This exemplified heavily by the interview's suggestion of the importance of empathizing and wanting to shadow the domain experts in their core services. The informants suggested automation to be a Big Data feature that project commissioners often find favorable. Thus, highlighting that concept of improving existing processes.

Moreover, Vanauer et al. (2015) suggest a second approach for ideating with Big Data, which is Data First. However, DF is mentioned by the informants as a bad practice as. Furthermore, it seemed that much of the data from the organization core services were more complex to obtain. It was either locked in other applications and therefore not accessible for Department A. One of the key features that some of the informants mentioned was the accessibility to data. The organization uses different software to overcome those problems by introducing platforms for storing data from different sources and computing power. However, as the organization is a service provider the data that they receive from the customer might not be as accessible for everyone in the department except from those assigned to the project. Thus, having a DF approach was therefore, not relevant to some extent.

Alam (2002) states that involving users in the Ideation stage is beneficial for product or service success. The finding from Department A shows that they have a similar approach for ideating. However, it becomes less evident when the project stems from external customers. In those scenarios Department A are entering the discussion quite late. Furthermore, informants state that ideating in terms of BDA becomes hard as other stakeholders have little knowledge of the field suggesting that occasionally, infeasible solutions are proposed. Magnusson (2009) states although lack of knowledge does provide an external dimension beyond the domain, some prior knowledge is beneficial. Magnusson (2009) states having guided users enables those users to ideate better as they have more understanding of the field. This coincides with the findings from the informants. Numerous cases were brought during the interviews where the interviewees stated that the discussion and Ideation regarding Big Data was hard with stakeholders with low BDA knowledge.

5.9 Big Data and Innovation

Big Data is mentioned by Loza et al. (2019) to be a co-innovation enabler. Stakeholders from different fields can cooperate by giving insights through their expertise to develop a solution. It was mentioned by the informants that this was the preferred way. Projects coming from inter-internal incentives were easier to initiate a discussion with the specific domain. Informants stated that most often, the workshop was often initiated with inter-domains rather than with external customers.

Moreover, Rindfleisch (2017) presented two ways of innovating, IFD (Innovation from Data) and IAD (Innovation As Data). As Department A is quite new, few products have emerged from them. Thus, Department A innovation from data is situated towards what data is received from intra-external stakeholders or external stakeholders. However, as Department A developed more products their chamber of data will grow enabling IFD to become more relevant.

6 CONCLUSIONS

This chapter presents and discusses the main findings by addressing the research questions. Recommendations for Department A will be discussed and suggestions for future research will be made.

RQ1: What challenges arise for the organization when ideating with Big Data?

1. Firstly, one of the challenges the organization has is not being Data First in their approach as the organization offers consultancy type of services. DF - Data First explains the approach where data is used as the primary source for developing value propositions. However, as the data is harder to access as it is clientele data and not the company's data being DF becomes more complex. Furthermore, being Data First also requires employees to have in-depth knowledge in a specific field but as customers could come from different industries it might be complex to withhold that knowledge. Thus, Data First is an approach less suitable for organizations working within a consultancy. The diversity of industries and the lack of data commissioner makes it problematic for companies within consultancy to be Data First. Hence, organizations in other industries would, therefore, have an easier task pursuing DF as data is easier to obtain and the domain-knowledge is present. BF - Business First is more suitable for consultancy companies as BF is initiated by understanding the business requirements and adding data where it is suitable.

2. Secondly, being BF results in less radical innovation to be developed. As new commodities are developed to evolve existing business processes less radical innovation is pursued. In order for the organization to be ideate more radical innovations data sources from clients must to be available not only when projects are initiated to enable DF. Thus, generally organization must allow data to be accessible for employees to explore in order for more radical innovation to prosper.

3. Thirdly, another challenge is communicating Big Data as it is a bit complex. Discussing theoretically about the opportunities surrounding Big Data might impair the discussion with stakeholders with less understanding. Conducting workshops is beneficial. However, a more streamlined and clear view of what Big Data is capable of could help other participants to be more constructive during Ideation. Thus, incorporating a data innovation board could be beneficial to reinforce the key components needed for utilizing Big Data.

RQ2: How can a better understanding of Ideation at an early stage in a project help organizations utilize the benefits of Big Data in a more effective way?

For the organization to utilize Big Data more effectively it is important that they consider:

1. Allowing Big Data experts to be more active with users during Ideation can help the organization to become more effective when finding opportunities with Big Data. Seemingly, introducing Big Data experts late into the Ideation discussion can hamper the innovation. The sourcing of different knowledge from several pipelines is important for innovation. Thus, ideating without Big Data experts can cause opportunities to be missed. Furthermore, an unfeasible discussion might be discussed causing the Big Data experts to redo much of the work again. Therefore, having a diverse group constellation consisting

of not only stakeholders with similar knowledge but other attributes such as a more in-depth understanding of Big Data is vital to use when suitable. Furthermore, the usage of design thinking often becomes harder to use productively if key contributors are not participating. Thus, integrating Big Data experts and the users in the Ideation stage early enables these experts to empathize much easier to the case. Moreover, as the organization is in audit, tax and advisory their customers are often top managers. Hence, lead-users, users with tremendous knowledge that can influence the development of more radical and functional innovation. Hence, by allowing Big Data experts to conjoin with other domains earlier in the Ideation stage with customers it will positively influence their product or service success.

2. Moreover, to be able to benchmark different ideas and thus, utilizing the Ambition Matrix as seen in Appendix B more concretely could help Big Data experts sustain the quality of the ideas and screen projects more efficiently. If not, data experts could miss out developing radical or disruptive innovation as the Ideation could be solely developed in the adjacent sphere of the Ambition Matrix. Thus, stalling the innovation degree of data experts.

The above suggestions are closely tied with upper management issues. Organizations must discuss which approach is most suitable for them BF or DF. What organizations should consider is in what sphere within the Ambition Matrix different projects should be in? Having a BF approach results in more incremental innovation whilst DF can spark more radical innovation. Thus, constructing clear innovation goals could contribute data experts to develop projects that are more in line with the corresponding goal of the organization. If the organization aims to pursue a more DF type of approach the organization should allow data experts to have the liberty to explore the data. In this case, considering the organization is a consultancy company should engage in projects that allow business customer data to be collected, thus building up a bank of obtainable data. Moreover, defining a clear innovation goal can help data experts to approach the Ideation, thus allowing the organization to have a more effective FEI.

6.1 Theoretical Implications

This thesis contributes to theory by linking the rising topic of Big Data with innovation and Ideation. Few studies have examined the Ideation stage with Big Data. Instead, most literature has explored the technical implication and the business opportunities organization could enable through Big Data. This research focuses on the Ideation stage of Big Data. Thus, identifying how a better understanding of Ideation makes the utilization of Big Data more effective. Moreover, the challenges of ideating with Big Data have also been presented in this thesis. Hence, highlighting the several aspects to consider when ideating.

6.2 Practical Implications

We propose some suggestions to consider when ideating with Big Data. The suggestions could be used to enable more effective utilization of Big Data. The analysis showed that there exposed factors that the organization and the different stakeholders should consider.

Firstly, Big Data is an enabler of co-innovation. Thus, it is important to consider the fact that Big Data does not generate value primarily because it is Big Data but that it can depict factors and

insights that might not have been considered before. However, the value that can be excavated from Big Data is dependent on the understanding of the specific field that is retrieved from. Thus, concatenating and combining other sources of Big Data becomes more useful if there is an underlying understanding of the domain. Thus, the Ideation of Big Data must be developed through joint creation with different stakeholders that allow a versatile combination of skills to approach the problems. Hence, Big Data experts must be more intertwined with the business unit meeting the external customers much earlier in the project process.

Secondly, to be able to ideate more proficiently all parties must have a fundamental understanding of Big Data and BDA. Otherwise, a rich discussion might be hard to reach as the different parties have difficulties understanding. It is therefore important that experts within Big Data have clearer ways of conveying the different aspects that must be considered before initiating a BDA type of approach. A similar Ideation board as the Data Innovation Board presented in section 2.6 could be adopted to help Big Data experts simplify Big Data into smaller and more understandable topics.

Thirdly, having a business-first approach to innovate with Big Data is a seemingly constructive approach. To become DF organizations and data experts must consider structuring up a predefined process of how internal data should be available. Otherwise, DF will be less and less relevant as data will become locked in a vast amount of applications that are developed. Moreover, defining what innovation is an aspect to consider. To be able to benchmark how the organization is performing becomes significantly more important in the future as the organization wants to see how they are performing. Neglecting this can increase the risk of organizations to continuously develop similar products repetitively.

6.3 Further Research

The findings from this research introduce a general aspect of how to ideate with Big Data in the early stage. The scope of this thesis was quite broad thus giving insights into the challenges of ideating with Big Data. Further research could be investigating more in-depth micro-level setting within Ideation of Big Data.

An area that could be examined is to see how much understanding of BDA is needed from the different parties to facilitate the richest discussion. Suggestion from the literature stated that having low understanding does not necessarily hamper innovation outcome but can launch ideas that might be unimaginable by the domain expert. Furthermore, some understanding is needed as otherwise, propositions might be unfeasible.

Moreover, another area that would be interesting to further investigate is how the Ideation process would look like when the problem is defined, and the different domain expertise is available. Thus, investigating how the different combination of Big Data could help the Ideation process for the organization.

6.4 Reflections

The case department which this thesis was carried out for is still quite new as it first was developed roughly two years ago. Thus, most of the employees in this group had not been in the organization for more than a year. Moreover, the department had few finished projects and was still setting up

its infrastructure in the firm. This caused some complications as few concrete experiences were available for the informants. Furthermore, examining Ideation with Big Data in a firm provides services within advisory, audit and tax to other organizations might not have been the best option in such a new topic. It could seemingly be more beneficial to investigate a similar department in an organization with an offering more intertwined with products that are less “sensitive”. Thus, the availability of actual data and the practical process within the organization was hard for us to participate in as the external stakeholders might only want to share it to relevant stakeholders. Hence, no dataset or actual solutions were presented to us limiting the investigation area to a much broader topic.

7 REFERENCES

Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., Gehrke, J., Haas, L., Halevy, A., Han, J. and Jagadish, H.V., 2011. Challenges and opportunities with Big Data 2011-1.

Ahuja, Gautam. "Collaboration Networks, Structural Holes and Innovation: a Longitudinal Study." *Academy of Management Proceedings*, vol. 1998, no. 1, 1998, pp. D1–D7., doi:10.5465/apbpp.1998.27664401.

Alam, I., 2002. An exploratory investigation of user involvement in new service development. Journal of the Academy of Marketing Science, 30(3), pp.250–261.

Alpaydın E. Introduction to Machine Learning. Adaptive Computation and Machine Learning Series. 2nd ed. Cambridge, MA: MIT Press; 2010.

Alpaydın, E., 2011. Machine learning. Wiley Interdisciplinary Reviews: Computational Statistics, 3(3), pp.195–203.

Bennett, R.C. and Cooper, R.G. (1979). Beyond the Marketing Concept. Business Horizons 22:76–83 (June).

Bennett, R.C. and Cooper, R.G. (1981). The Misuse of Marketing: An American Tragedy. Business Horizons 24(6):51–61.

Bharadwaj, N., and C. H. Noble. 2015. Innovation in data-rich environments. Journal of Product Innovation Management 32: 476–78

Bogers, Marcel, et al. "The 'Human Side' of Open Innovation: The Role of Employee Diversity in Firm-Level Openness." *Research Policy*, vol. 47, no. 1, 2018, pp. 218–231., doi: 10.1016/j.respol.2017.10.012.

Brown, Tim, and Barry Katz. "Change by Design." *The Journal of Product Innovation Management*, vol. 28, no. 3, 2011, p. 381., doi:10.1111/j.1540-5885.2011.00806. x.

Carr, L.T. (1994), The strengths and weaknesses of quantitative and qualitative research: what method for nursing? Journal of Advanced Nursing, 20: 716-721. doi:10.1046/j.1365-2648.1994.20040716.x

Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. Mobile Networks and Applications, 19(2), 171-209.

Cisco (2016). "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020." *VNI Global Fixed and Mobile Internet Traffic Forecasts*, Cisco, 3 Feb. 2016, www.cisco.com/c/en_ca/solutions/service-provider/visual-networking-index-vni/index.html

Court, David. “Getting Big Impact from Big Data.” *McKinsey & Company*, Jan. 2015, www.mckinsey.com/business-functions/mckinsey-digital/our-insights/getting-big-impact-from-big-data.

Coyle, K. (2006), “Mass digitization of books”, Journal of Academic Librarianship, Vol. 32 No. 6, pp. 641-645

Crossan, Mary M., and Marina Apaydin. “A Multi-Dimensional Framework of Organizational Innovation: A Systematic Review of the Literature.” *Journal of Management Studies*, vol. 47, no. 6, 2009, pp. 1154–1191., doi:10.1111/j.1467-6486.2009.00880. x.

Dahlander, Linus, and David M Gann. “How Open Is Innovation?” *Research Policy*, vol. 39, no. 6, 2010, pp. 699–709., doi: 10.1016/j.respol.2010.01.013.

“Datainspektionen.” *Datainspektionen*, www.datainspektionen.se/lagar--regler/dataskyddsforordningen/grundlaggande-principer/.

Davenport, T. (2013). Harvard Business Review, Harvard business review.

Davenport, T. & Patil, D., 2012. Data Scientist: The Sexiest Job of the 21st Century. Harvard Business Review, 90(10), pp.70–76.

Devers, Kelly J, and Richard M Frankel. “Study Design in Qualitative Research-2: Sampling and Data Collection Strategies.” *Education for Health: Change in Learning & Practice*, vol. 13, no. 2, 2000, pp. 263–271., doi:10.1080/13576280050074543.

Dewar, Robert, D., and Jane E. Dutton. “The Adoption of Radical and Incremental Innovations: An Empirical Analysis.” *Management Science*, vol. 32, no. 11, 1986, pp. 1422–1433., doi:10.1287/mnsc.32.11.1422.

De Mauro, A., Greco, M. & Grimaldi, M., 2016. A formal definition of Big Data based on its essential features. Library Review, 65(3), pp.122–135.

Dodge, Richard, et al. “The Role of Leadership in Innovation: A Quantitative Analysis of a Large Data Set Examines the Relationship between Organizational Culture, Leadership Behaviors, and Innovativeness.” *Research-Technology Management*, vol. 60, no. 3, 2017, pp. 22–29., doi:10.1080/08956308.2017.1301000.

Duffy M E (1986) Quantitative and qualitative research antagonistic or complementary' Nursing and Health Care 8(6), 356-357

Eling, Katrin, and Cornelius Herstatt. “Managing the Front End of Innovation—Less Fuzzy, Yet Still Not Fully Understood.” *Journal of Product Innovation Management*, vol. 34, no. 6, 2017, pp. 864–874., doi:10.1111/jpim.12415.

Eisenhardt, K.M. (1989) “Building theories from case study research.” Academy of Management Review, 14(4), pp.532-550.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. Journal of Business Research, 69(2), 897-904.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big Data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137-144.

Garcia, Rosanna, and Roger Calantone. "A Critical Look at Technological Innovation Typology and Innovativeness Terminology: a Literature Review." *The Journal of Product Innovation Management*, vol. 19, no. 2, 2002, pp. 110–132., doi:10.1016/S0737-6782(01)00132-1.

GDPR, Principles relating to processing of personal data. Art 5 Financial Assistance to Individuals, 45 C.F.R. § 234 (2012).7 C.F.R 1 319 (2000).

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. International Journal of Market Research, 59(6), 703-705.

Graham, Douglas, and Thomas T.Bachmann. *Ideation: the Birth and Death of Ideas.* John Wiley and Sons, 2004

.

Gruschka, N. et al., 2018. Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR.

IBM, "IBM Big Data: A collection of Big Data client success stories," 2012.

IDC (2015) "Revenues for Big Data and Business Analytics Solutions Forecast to Reach $260 Billion in 2022, Led by the Banking and Manufacturing Industries, According to IDC." *IDC*, IDC, Aug. 2015, www.idc.com/getdoc.jsp?containerId=prUS44215218.

Ishwarappa, J. & Anuradha, 2015. A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology. Procedia Computer Science, 48(C), pp.319–324.

ITGP Governance Privacy Team, (2016). EU General Data Protection Regulation (GDPR) : An implementation and compliance guide.

Kakatkar, C., Bilgram, V. & Füller, J., 2019. Innovation analytics: Leveraging artificial intelligence in the innovation process. Business Horizons, 63(2), pp.171–181.

Khurana, Anil, and Stephen Rosenthal. "Towards Holistic 'Front Ends' in New Product Development." *The Journal of Product Innovation Management*, vol. 15, no. 1, 1998, pp. 57–74., doi:10.1016/S0737-6782(97)00066-0.

Kirk, J. & Miller, M., 1986. Reliability and Validity in Qualitative Research, Newbury Park: SAGE Publications, Inc.

Kiron, David. "Lessons from Becoming a Data-Driven Organization." *MIT Sloan Management Review*, 18 Oct. 2016, sloanreview.mit.edu/case-study/lessons-from-becoming-a-data-driven-organization/.

Koen, Peter, et al. “Providing Clarity and A Common Language to the ‘Fuzzy Front End.’” *Research-Technology Management*, vol. 44, no. 2, 2001, pp. 46–55., doi:10.1080/08956308.2001.11671418.

Kristensson, P., Gustafsson, A. & Archer, T., 2004. Harnessing the Creative Potential among Users *. Journal of Product Innovation Management, 21(1), pp.4–14.

Kronsbein, Tizian, and Rolland Mueller. “Data Thinking: A Canvas for Data-Driven Ideation Workshops.” *Scholar Space at University of Hawaii at Manoa: Home*, Hawaii International Conference on System Sciences, 8 Jan. 2019, scholarspace.manoa.hawaii.edu/handle/10125/59496.

Lapan, S.D., Quartaroli, M.T. & Riemer, F.J., 2012. Qualitative research an introduction to methods and designs First., San Francisco: Jossey-Bass.

Lau, R., Zhao, J., Chen, G., & Guo, X. (2016). Big Data commerce. Information & Management, 53(8), 929-933.

Lee, S., Olson, D., & Trimi, S. (2012). Co-innovation: Convergenomics, collaboration, and co-creation for organizational values. Management Decision, 50(5), 817-831.

Li, H., Yu, L. & He, W., 2019. The Impact of GDPR on Global Technology Development. Journal of Global Information Technology Management, 22(1), pp.1–6.

Lozada, N., Arias-Pérez, J., & Perdomo-Charry, G. (2019). Big Data analytics capability and co-innovation: An empirical study. Heliyon, 5(10), E02541.

Magnusson, P., 2009. The journal of product innovation management, pp.578–593.

Maguire, M.S., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars.

McAfee, A. & Brynjolfsson, E., 2012. Big Data: The Management Revolution. Harvard Business Review, 90(10), pp.61–68.

Morrison, P.D, J.H., and von Hippel, E. (2000). Determinants of User Innovation and Innovation Sharing in a Local Market. Management Science 46(12):1513–27

Nagji, Bansi, and Geoff Tuff. “Managing Your Innovation Portfolio.” *Harvard Business Review*, 19 May 2015, hbr.org/2012/05/managing-your-innovation-portfolio.

Norman, Donald A., and Robert Verganti. “Incremental and Radical Innovation: Design Research vs. Technology and Meaning Change.” *Design Issues*, vol. 30, no. 1, 2014, pp. 78–96., doi:10.1162/DESI_a_00250.

Prescott, A. (2013), “Bibliographic records as humanities Big Data”, Proceedings – 2013 IEEE International Conference on Big Data, IEEE, pp. 55-58, available at: http://doi.org/10.1109/ BigData.,2013.6691670

Pu Wang, Timothy Hunter, Alexandre M. Bayen, Katja Schechtner, & Marta C. González. (2012). Understanding Road Usage Patterns in Urban Areas. Scientific Reports, 2(1), 1001.

Raghupathi, W., Raghupathi, V. Big Data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2, 3 (2014). https://doi.org/10.1186/2047-2501-2-3

Ries, E. (2011). *The Lean Startup*. New York: Crown Business.

Rindfleisch, A., O’Hern, M. and Sachdev, V. (2017), “The digital revolution, 3D printing, and innovation as data”, Journal of Product Innovation Management, Vol. 34 No. 5, pp. 681-690.

Russom, Philip (2011). *Big Data Analytics*. Retrieved from TWDI Best practices report https://nccam.nih.gov/sites/nccam.nih.gov/files/news/nhsr12.pdf

Sagiroglu, S. & Sinanc, D., 2013. Big Data: A review. 2013 International Conference on Collaboration Technologies and Systems (CTS), pp.42–47.

Sandelowski M (1986) The problem of rigor in qualitative research Advances m Nursing Science 8(3), 27-37

SAP. (2014). Real-time enterprise stories. Retrieved October 9, 2016, from http://www.sap.com/bin/ sapcom/en_us/downloadasset.2014-10-oct-13-21.sap-hana-real-time-enterprise-stories-pdf. bypassReg.html. +

Shvachko, K., Kuang, H., Radia, S. and Chansler, R. (2010), “The Hadoop distributed file system”,IEEE 26th Symposium on Mass Storage Systems and Technologies, MSST2010, IEEE, pp. 1-10.

Strauss, A., Corbin, J. 1998. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications.

Sumbal, M., Tsui, E., Irfan, I., Shujahat, M., Mosconi, E., & Ali, M. (2019). Value creation through Big Data application process management: The case of the oil and gas industry. Journal of Knowledge Management, 23(8), 1566-1585.

Quintane, Eric, et al. “Innovation as a Knowledge-Based Outcome.” *Journal of Knowledge Management*, vol. 15, no. 6, 2011, pp. 928–947., doi:10.1108/13673271111179299.

Tankard, C. (2016). What the GDPR means for businesses. Network Security, 2016(6), 5-8.

Tivārī, R. (2015). Hadoop for finance essentials: Harness Big Data to provide meaningful insights, analytics, and business intelligence for your financial institution (Community experience distilled).

Tongco, Ma. Dolores C. “Purposive Sampling as a Tool for Informant Selection.” *Ethnobotany Research and Applications*, vol. 5, 2007, p. 147., doi:10.17348/era.5.0.147-158.

M. Vanauer, C. Böhle and B. Hellingrath, "Guiding the Introduction of Big Data in Organizations: A Methodology with Business- and Data-Driven Ideation and Enterprise Architecture Management-Based Implementation," *2015 48th Hawaii International Conference on System Sciences*, Kauai, HI, 2015, pp. 908-917. doi:10.1109/HICSS.2015.113

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.

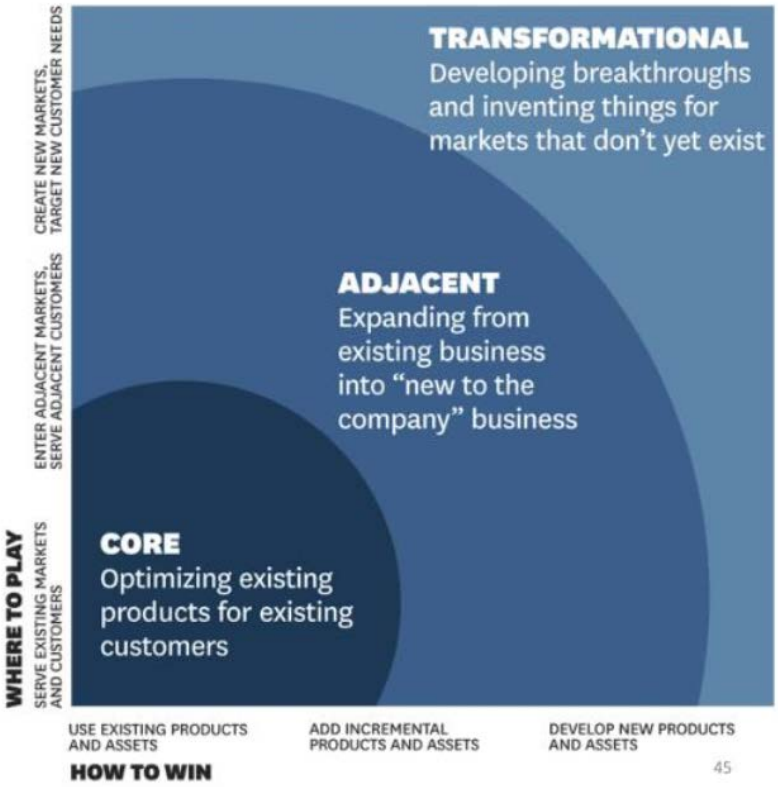
Ward, J.S. & Barker, A., 2013. Undefined By Data: A Survey of Big Data Definitions.

Zhan, Y., Tan, K., Ji, G., Chung, L., & Tseng, M. (2017). A Big Data framework for facilitating product innovation processes. Business Process Management Journal, 23(3), 518–536.

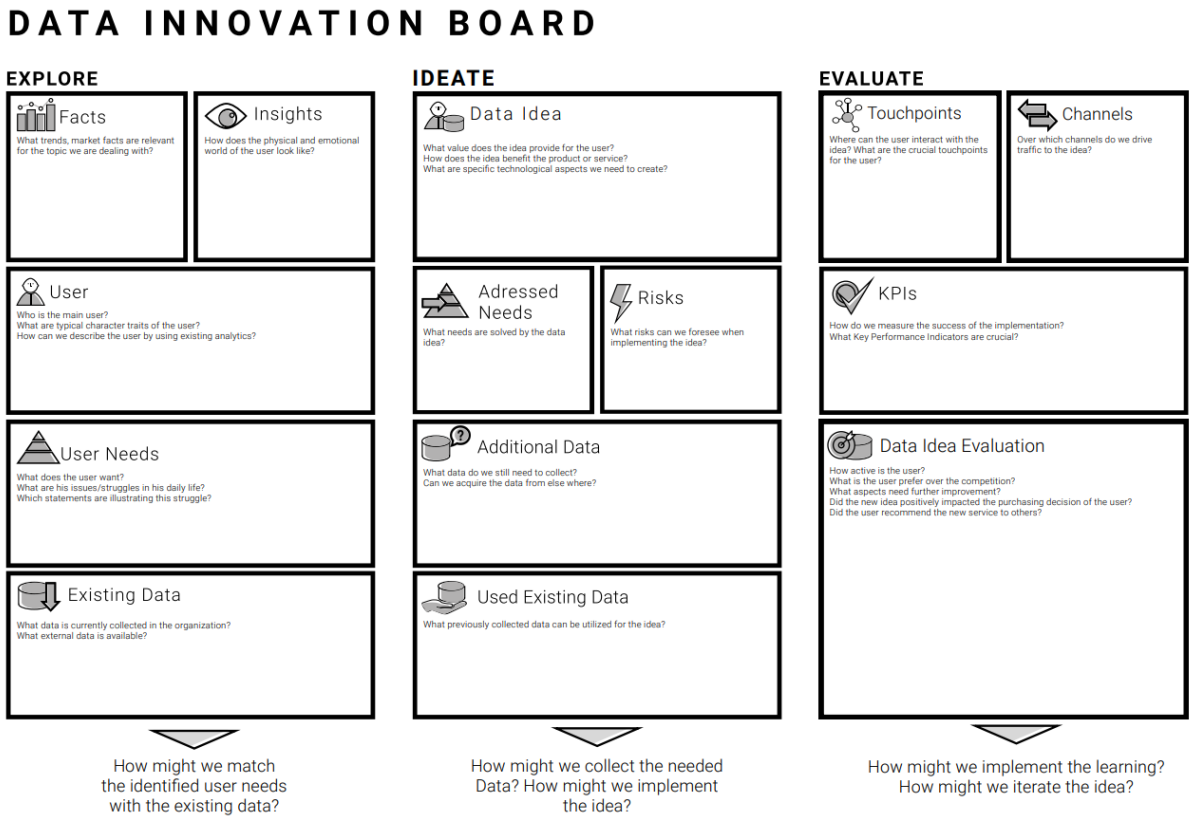
APPENDIX A: TABLE 1

Level	Role	Type
1	Head of Department A	Interview 45 min
1	Head of an External Department	Interview 45 min
2	Manager	Interview 45 min
2	Manager	Interview 45 min
2	Manager	Interview 45 min
2	Manager	Interview 45 min
2	Manager	Interview 45 min
3	Associate	Interview 45 min
3	Associate	Interview 45 min
3	Associate	Interview 45 min
3	Associate	Interview 45 min
3	Associate	Interview 45 min
3	Associate	Interview 45 min

APPENDIX B: AMBITION MATRIX



APPENDIX C: DATA INNOVATION BOARD



APPENDIX D: Interview Guide II

General questions about the company and interviewee

What is your role in the company?

Q1: Project

How do you get your projects? Does it depend on the type of project (internal/external)?

Are you in direct contact with the end-customers if the project is external right from the start? Or do you come in after the project has been initiated?

Where in the process of a project are you coming into the picture?

Which people are responsible for meeting the other department when a project is proposed (internally/externally)?

Q2: General

How does the first phase within idea/need generation look like for you? Do you have any examples?

Lead Generation?

What is Big Data Analytics and how has it influenced FEI for you?

How is data exploration for you? Do you do it individually or in teams? Could it be improved by collaborating with different experts in the specific domains you are working with?

Q3: Ideation

Does it often happen that the problem is not well defined from the other departments?

Would it make a big difference on the need generation and problem definition if Lighthouse could attend the first customer meetings together with other departments?

What type of tools do you use for defining a problem when the customers are not sure about the problems that they have? Storyboarding, mind, mapping, flowcharts?

What tools do you use for generating possible ideas with the customers? Brainstorming, creative collaboration? Do you use any tools for looking at the problem from an obscure angle? Scamper, triggering?

How would you want to work with problem definition and need generation?

Are workshops a method you use often? Could you define a workshop?

When do you have workshops at Insight Center? Who is deciding to have them?

Lighthouse strategy is, innovate with, deliver with and sell with. How much do you innovate with if you are not part of the first initial lead generation meeting?

What is innovation for you within your working processes? Do you benchmark different projects with the grade of innovation? What is core and what is radical for you?

Would you consider most of the projects that you deliver to be innovative?

Q4: Big Data and Business opportunities

How has your expertise in data analytics made new business opportunities? Do you have any examples?

Are there times when you have drawn insights from data that your client has not thought of? skippa

Have your expertise created new opportunities that your other colleagues (external to lighthouse) have not been able to define?

What are the key features Big Data brings to creating business opportunities?

TRITA TRITA-ITM-EX 2020:405