



LUNDS UNIVERSITET

Ekonomihögskolan

Department of Informatics

Big Data IT Capabilities

Bachelor Thesis 15 hp, Course SYSK02 in information systems

Authors: Mattias Rasch Skogh
Viktor Karlsson

Supervisor: Styliani Zafeiropoulou

Examiners: Björn Johansson
Anders Svensson

Big Data IT Capabilities

Authors: Mattias Rasch Skogh and Viktor Karlsson

Publisher: Dept. of informatics, School of economics and management, Lund University

Document type: Bachelor thesis

Pages: 73

Key words: Big data, RBV, IT Capabilities, Value creation

Abstract

The amount of data produced is growing rapidly and organizations are struggling to make use of the generated data from the vast amount of sources. Analyzing and interpreting data to find patterns and trends has become a competitive advantage for companies working with customer oriented operations. Many previous studies have shown the value potential of working with big data. In order to create value from big data an increased demand on companies IT capabilities has risen. We asked ourselves the question, which IT capabilities are these? To evaluate the capabilities needed, we have applied the Resource Based View (RBV). The RBV addresses the complementarity of IT and organizational processes and practices in organizational resources. To answer our question, we examined academic literature regarding big data and RBV and put together a number of discussion points, which we then used in interviews with five customer oriented companies. After conducting the interviews an analysis was performed. The analysis compares the literature to the empirical data and commonalities and discrepancies were identified. We found that strong teams with several different competencies are required when working with big data. Furthermore, a new culture is arising with big data. A culture that requires decision makers to make decisions based on data instead of intuition. For the future, we recommend all managers to learn how big data can create value for their organization.

Table of contents

| | |
|--|----------|
| Big Data IT Capabilities | 2 |
| Abstract | 2 |
| 1. Introduction | 6 |
| 1.1 Background | 6 |
| 1.2 Problem discussion | 7 |
| 1.3 Purpose and research question | 8 |
| 1.4 Delimitation | 8 |
| 2. Literature Review | 8 |
| 2.1 Literature discussion | 9 |
| 2.2 Previous research..... | 10 |
| 2.3 Big data concept | 11 |
| 2.4 Big data process | 12 |
| 2.5 Big data in the corporate context | 13 |
| 2.5.1 How big data is changing sales and marketing | 13 |
| 2.5.2 Big data management..... | 14 |
| 2.5.3 Big data applications..... | 14 |
| 2.6 Big data benefits and opportunities | 15 |
| 2.7 Risks and challenges of big data | 15 |
| 2.7.1 Issue areas..... | 16 |
| 2.8 Requirements for success with big data | 17 |
| 2.8.1 Big data strategy | 17 |
| 2.8.2 Key terms for success..... | 18 |
| 2.9 Big data sources | 18 |
| 2.10 Other use-areas of big data | 19 |
| 2.11 The Resource Based View | 20 |
| 2.11.1 Resources | 21 |
| 2.11.2 Assets | 21 |
| 2.11.3 Capabilities..... | 21 |
| 2.11.4 IT Resources | 22 |
| 2.11.5 IT Assets | 23 |
| 2.11.6 IT Capabilities | 23 |

| | |
|--|-----------|
| 2.12 Literature Summary | 25 |
| 3. Method..... | 26 |
| 3.1 Data collection method..... | 26 |
| 3.2 Company and participant selection | 26 |
| 3.2.1 Participants | 27 |
| 3.3 Conducting the interviews | 28 |
| 3.3.1 Interview guide | 28 |
| 3.4 Data Analysis | 29 |
| 3.5 Research Quality | 30 |
| 3.5.1 Validity | 30 |
| 3.5.2 Reliability..... | 30 |
| 3.5.3 Ethics..... | 31 |
| 4. Results..... | 32 |
| 4.1 IT capabilities | 32 |
| 4.1.1 Competencies | 32 |
| 4.1.2 Practices | 34 |
| 4.2 Value creation | 35 |
| 4.2.1 How big data creates value for organizations | 35 |
| 4.2.2 Value opportunities | 36 |
| 4.2.3 Big data and marketing..... | 36 |
| 5. Discussion..... | 37 |
| 5.1 IT Capabilities | 37 |
| 5.1.1 Competencies | 37 |
| 5.1.2 Practices | 37 |
| 5.3 Value creation | 38 |
| 6. Conclusion | 40 |
| 7. Appendix | 41 |
| Appendix 1 - Interview Guide | 41 |
| Appendix 2 – Styleport..... | 42 |
| Appendix 3 – Softhouse | 46 |
| Appendix 4 – Lundalogik..... | 53 |
| Appendix 5 – Avensia | 60 |
| Appendix 6 – King | 65 |
| References..... | 70 |

Figures

| | |
|--|----|
| Figure 1, 5 Vs of Big Data (Demchenko, Grosso, Laat & Membrey, 2013) | 11 |
| Figure 2, The big data process (Saravana, 2013) | 13 |
| Figure 3, Theoretical Model of IT Resources (Aral and Weill, 2007) | 22 |
| Figure 4, Theoretical Model used for our research. Based on Theoretical Model of IT Resources (Aral and Weill, 2007) | 24 |
| Figure 5, The competencies derived from the theoretical model | 32 |
| Figure 6, The practices derived from the theoretical model | 34 |

Tables

| | |
|---|----|
| Table 1, Previous literature covering the Resource Based View, Big data and IT Capabilities | 10 |
| Table 2, Companies interviewed | 28 |

1. Introduction

In this initial chapter the background, problem, research question as well as thesis and delimitation are introduced.

1.1 Background

From the early forms of communication to the cloud technology of today, humans have been gathering information. The last century's rapid development of technology has led to an overflow of data, with increasing demands on the data storage systems. As early as 1944, Fremont Rider, a Wesleyan University Librarian, publishes *The Scholar and the Future of the Research Library*. In the book the librarian flagged a warning of the future challenges of handling data. He estimated that American libraries doubles in size every sixteen year. With this growth rate the Yale Library would have approximately 200,000,000 books over 10,000 kilometers of shelves in the year of 2040. The librarian could not foresee how the upcoming digital technologies would take over as the primary source of data storage.

During the Digital Revolution in the 1960s and 1970s mechanic and electronic technology changed to digital technology. Adoption of digital computers began and increasing amounts of data started to be stored through digital records. The society became more information oriented and the information era began. The increased use of computers and the rapid technology development led to a massive exponential growth of produced data. In the year of 2000, Peter Lyman and Hal R. Varian publishes the book "*How much Information?*" The book is the first comprehensive study which quantifies the total amount of data created in computer storage terms. The study also finds that the digital data produced is not only largest in total but also the most rapidly growing (Press, 2013).

In 2005 the term 'Big data' was coined by Roger Magoulas from O'Reilly media. The term refers to digitally stored data and datasets of such a complexity or extensive size that make traditional data processing techniques inadequate (Halevi & Moed, 2012).

The digitalization of data has been growing rapidly for the last decades but it is only over the last few years that big data has emerged as the new frontier in the wide spectrum of IT enabled innovations and opportunities. This is allowed by the fast development of information technology (Goes, 2014). According to an IBM infographic the world is producing 2.5 quintillion (2.5×10^{18}) bytes of data every day through data generating sources like social media, internet of things and wearable monitors (IBM, 2013). According to Gartner a total of 5 billion things were connected to the internet in 2015, constantly feeding data to organizations and cities. An infographic published in 2015 by Aureus Analytics states that data is growing by 40% annually and that it will grow 50 times by 2020, the data explosion is not about to slow down (Aureus Analytics, 2015).

The vast amount of data created has led to the big question of what to do with the data. How do you harness all the data and create value from it. Different actors see different uses of big data; the organizations are looking for solutions to use big data to increase their competitive advantage, while academia and scientific sectors tries to find new and better ways to understand the world (Goes, 2014).

1.2 Problem discussion

Companies analyzing these large chunks of data helps them identify trends and user behavior to enhance their business decisions. But it also puts demands on digital competence in order to reap the full benefits of big data. A recent study made by Fujitsu shows that 52% of Swedish decision makers consider their digital strategy obscure and confusing (Ekelund, 2016). Many Swedish companies see themselves as in the front edge of digitalization but the study shows the complete opposite in comparison with their European counterparts. The comparison also shows that the Swedes lack both in competence and engagement.

Karin Schreil Jonsson, CEO of Fujitsu Sweden who is behind the study claims *“We have both the infrastructure and the networks that enables a digital way of work but we lack the knowledge to take advantage of the technique.”* (Ekelund, 2016 p. 1). The study also shows that 3 out of 4 feel that the digital projects they take on are gambles whether the project will be successful or not.

Managers and board members among the Swedish stockbroking companies have a hard time keeping up in this digital era. A study conducted by Dagens industri shows that 8 out of 10 board members in the OMXS30-companies are missing digital competence (Carlsson, 2016). *“Swedish companies are really far behind”* claims digitalization expert Sara Öhrvall. *“It’s hard to determine how much a new technique is able to change. There have been too many examples of companies who repeatedly have miscalculated this.”* (Carlsson, 2016, p. 1). Big company names like SEB, Lundin Petroleum, Alfa laval, Atlas Copco and Skanska have boards lacking digital background entirely. With the lack of digital competence, problems may arise for several reasons. The biggest issue is that the board will ask management the wrong questions. *“The focus is not placed on the right questions to discuss.”* The lack of digital competence might also mean missed opportunities for streamlining processes (Carlsson, 2016, p. 1).

A report by the McKinsey Global Institute (Chen, Chiang and Storey 2012) predicted that by 2018, the United States alone will have a shortage of 140,000 to 190,000 people with deep analytical skills. Furthermore there will be a lack of managers with the relevant digital competency of utilizing big data in order to gain insights and to make effective decisions accordingly. The statement from Hal Varian, Chief Economist at Google and emeritus professor at the University of California, Berkeley, tells us his view on the future development of big data: *“So what’s getting ubiquitous and cheap? Data. And what is complementary to data? Analysis. So my recommendation is to take lots of courses about how to manipulate and analyze data: databases, machine learning, econometrics, statistics, visualization, and so on.”* (Chen et al., 2012 p. 2).

Facing the need to analyze and interpret data has resulted in a brand new industry. Uber and Airbnb are two examples of companies that effectively uses big data to empower their businesses and they are both claiming huge market shares. Both Uber and Airbnb shows that digitization is overturning many industries and an increased need for businesses to adapt to the changing environment in order to stay competitive. Managers need to realize this and form their strategy and behavior accordingly. The amount of collectible data is growing rapidly and companies able to utilize this big data are the winners at the end of the day. The idea is to better understand the environment, both the external and internal through the analyzation of big data, which we will

show by proposing business examples. By understanding both the arising customer trends and the internal work processes of the organization, the company can obtain competitive advantage.

We have to understand that the collection of big data is not a solution in itself, it is only a tool. Correct usage of the tool, meaning thorough analysis, trend identification and lastly for decision makers to use these results that enable value creation (Watson, 2014).

1.3 Purpose and research question

The thesis aims to investigate which IT capabilities affect value creation from big data. We try to achieve a big picture of what IT Capabilities are needed and how organizations need to utilize them in order to create value. We start by interviewing different stakeholders, actively working with big data within companies with customer oriented operations. Thereafter we will compare our findings with past literature relevant to our subject. Based on these comparisons, we will analyze similarities and differences and draw a conclusion.

Research Question: *What IT capabilities does a company need in order to create value from big data?*

1.4 Delimitation

The study is limited to investigating what value big data can bring to organizations. To break down the subject into a more graspable scope we have chosen to look at the value big data can bring from a Resource Based View (RBV). To understand what IT Capabilities are, we explain what IT resources and IT assets are in regards to value creation of big data. We have narrowed it down further to primarily look at the IT capabilities necessary for organizations to utilize big data. However we still want to make it clear that IT capabilities are only a small piece of the puzzle of a successful company.

2. Literature Review

In this chapter we will treat the existing and relevant literature which matches our research question. Furthermore a theoretical basis for the discussion and the conclusion is created. Initially this chapter will provide a background for big data, its definition and its history. This is done in order to provide an understanding of how big data can create value for businesses and organizations.

We will also provide a background for the theories we are using throughout the thesis. The theories provide a basic understanding of what IT Capabilities are and important factors regarding value creation. In order to understand why we have chosen these theories we will present relevant articles and previous research. It is important to understand what big data is and how it can be used to create value for an organization.

2.1 Literature discussion

Since we focus on IT Capabilities and value creation we have chosen literature we find relevant in order to understand the Resource Based View as well as the research leading up to our final definition of IT Capabilities. The research is also used as a basis for understanding what big data is, the definition that we have chosen to apply in our thesis as well as some business examples. We have chosen not to include all literature in our table, only the articles we feel are related directly to our topic.

2.2 Previous research

| Author | Method | Context | Short Description | Theory |
|---|---|------------------------------|--|---|
| S. Aral, P. Weill (2007) | Survey | Employees/ Organizations | Creates a theoretical model for IT Resources out of a resource based view. | Resource-Based View of the firm, Resource-based Theory of IT. |
| N. Bekmamedova, G. Shanks (2014) | Case Study | Organization | Presents a theoretical framework that explains how organizations create value with SMA. A number of important lessons learned about effective use of SMA are discussed. | Organizational Motivation Theory, Resource Based View. |
| H. Chen, R.H.L. Chiang, V.C. Storey (2012) | Analysis of previous research | Organization | Provides a framework that identifies the evolution, applications and emerging research areas of Business Intelligence and analytics | N/A |
| J. Gao, S. Selle, A. Koronios (2015) | Content Analysis | Organizational | Provides critical success factors for the success of a big data project. In addition the thesis provides a process model as a suggestion of how to approach big data projects. | N/A |
| R. Santhanam, E. Hartono (2003) | N/A | Analytical | Evaluates the Resource Based View framework. Further examines firms' performance based on their IT capabilities. | Resource Based View |
| Y. Demchenko, P. Grosso, C. Laat, P. Membrey (2013) | N/A | Analytical | Discusses the definition of big data. Defines requirements on security, data management and access control. Also introduces the Scientific Data Life Cycle management model. | Scientific Data Infrastructure, Scientific Data Life Cycle Management |
| M. Wade, J. Hulland (2004) | Evaluation of the Resource Based View framework | Analytical | Explore and critically evaluate the use of Resource Based View. | Resource Based View |
| Hugh J. Watson (2014) | Literature review | Analytical | Big data analytics: concepts, technologies and applications. | N/A |
| D. Cole, J. Nelson, B. McDaniel (2015) | Literature review | General overview of big data | The paper explores big data, its previous and future uses and risks and benefits with it. | N/A |

Table 1. Previous literature covering the Resource Based View, Big data and IT Capabilities

2.3 Big data concept

There are many different definitions of big data. Sorofman, (2013, p. 1) describes big data as "*a cute way of describing the idea of data processed at massive scale and speed, where the trail thrown off by all of our varied digital interactions and experiences becomes the fuel for decisions, insights and actions*".

"*Big data is a concept that means, firstly, the volume of the data is too large. Secondly, it is impossible to analyze it using conventional technologies, and thirdly, special tools and treatment are necessary to extract knowledge from it*" (Manyika et al., 2011, p. 16).

"*Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making*" (Sicular, 2013, p. 1).

The definition used for this research is derived from Forbes author Lisa Arthur (2013, p. 1): "*Big data is a collection of data from traditional and digital sources inside and outside your company that represents a source for ongoing discovery and analysis.*"

This definition encompasses both the digital inputs like interactions on social media and web behavior as well as the more traditional data from transactions and financial records. To further describe and understand what big data is, it has been divided into five different aspects, five Vs: Volume, Velocity, Variety, Veracity and Value (Marr, 2014).

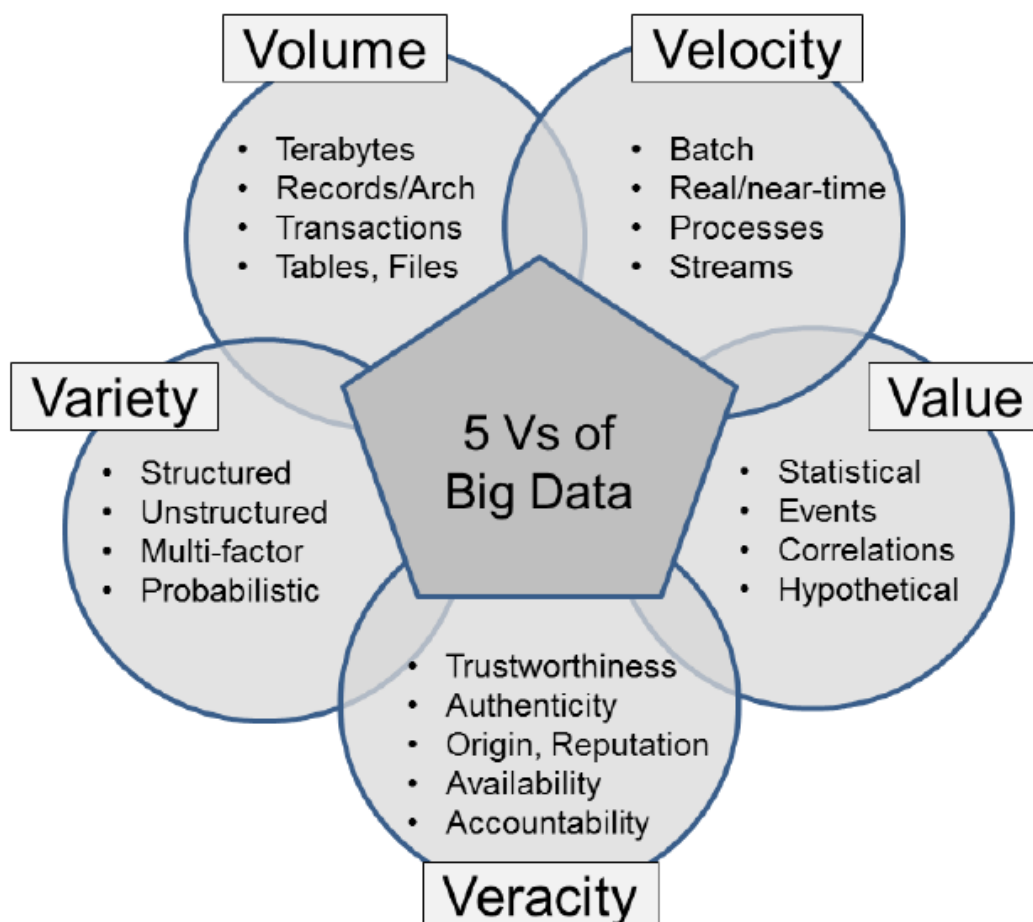


Figure 1, 5 Vs of Big Data (Demchenko, Grosso, Laat & Membrey, 2013)

Volume: Refers to all the data that is constantly generated. According to SINTEF (2013), 90% of the world's data has been produced in the last two years. The increase of produced data makes data sets too large to store or analyze using traditional data processing techniques (Marr, 2014). The new big data technologies are capable of handling the huge chunks of data through distributed systems which stores the data in different locations and then brings it together by using software.

Velocity: Refers to how fast the data is generated, processed and analyzed over time. It could for instance be the speed of which a credit card transaction is checked for fraud activities. The technology today allows companies to analyze the data without ever putting it into databases.

Variety: Refers to how many different data types we can use today. There has previously been a big focus on structured data, which is information that resides in fixed fields within a record or file. Usually this is data which is contained in spreadsheets and relational databases. Organizations are familiar with this type of data and many are using it effectively already (Marr, 2014). Marr further states that over 80% of business-relevant data is unstructured, which means that the data does not have an existing internal structure and cannot be put into tables. With big data technology it is now possible to use more types of data, including internet of things, sensors, social media, log files and messages which can bring unstructured data together with the structured data.

Veracity: Refers to how reliable the data is. With the many different types of big data it is hard to validate the data, sort out the noise and malicious data from the valid. An IBM infographic states that one in three managers does not trust the information they use to make decisions (IBM, 2013). The same infographic states that poor data quality costs the US \$3.1 trillion a year. Big data and analytics technology have become increasingly important for the improvement of data accuracy (Marr, 2014).

Value: The last aspect refers to what value can be extracted from the data. It all comes down to doing a cost and benefits analysis of what big data initiatives actually will bring to the company. Big data is useless if a company cannot turn it into value (Marr, 2014).

It is important to not see the different aspects as separate entities but instead combine these five dimensions. Integrating the data sources, being able to work with different data types, validating the data and to learn how to bring value out of it.

2.4 Big data process

In order to gain a better understanding of what big data is, we have chosen to include a model presenting the big data process.

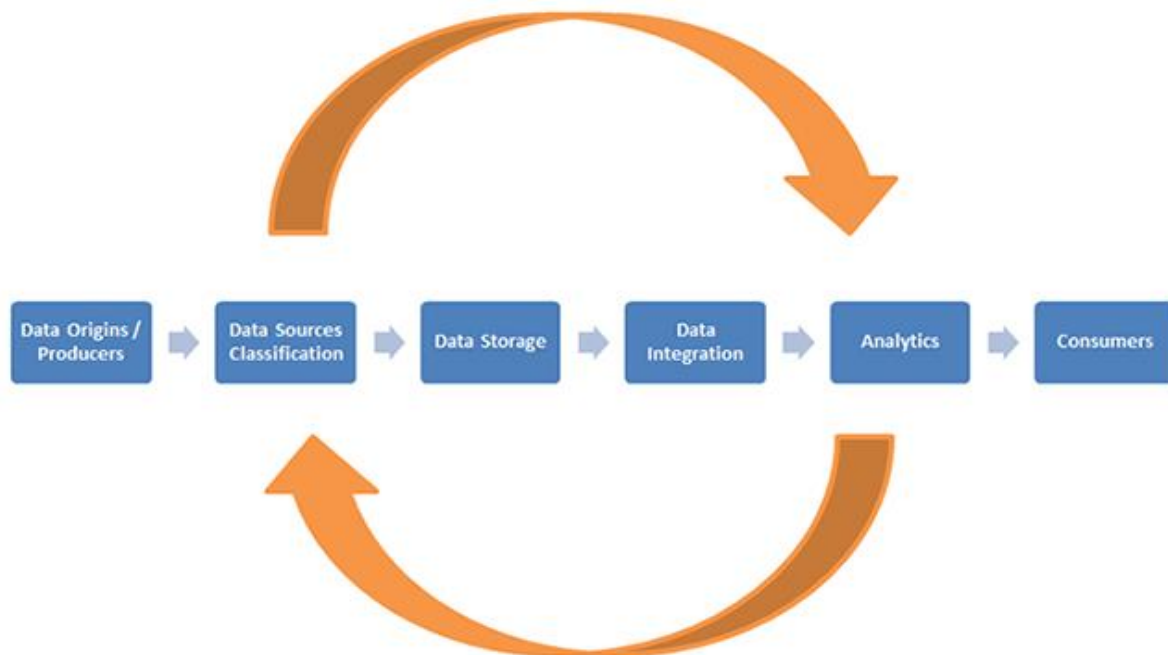


Figure 2, The big data process (Saravana, 2013)

The feedback loop and cycle are from data source classification to analytics:

Data Origins / Producers covers the Internet, sensors, machines and other types of origins. **Data Sources** are web logs, sensor data, images/audio, videos and more. **Data Storage** refers to Technologies supporting data storage. **Data integration** is the combination of technical and business processes used to combine data from disparate sources into meaningful and valuable information. **Analytics** means business analytics functions in regards to big data analysis. For instance, analytics can help find patterns in data to help decision makers or improve business processes. **Consumers** are customers, business processes and applications (Saravana, 2013).

2.5 Big data in the corporate context

There is significant interest in big data from a vast number of fields. The big data era is here and companies need to adapt to survive (Kiron, Ferguson and Prentice, 2013). We have chosen the corporate context, however there are other applicable fields which we are mentioning briefly.

2.5.1 How big data is changing sales and marketing

The uprising of big quantities of data is generating big challenges as well as opportunities for a large variety of companies. Hollison (2015, p. 1) claims that “*big data is the most transformative paradigm to hit sales and marketing teams since the advent of the telephone.*” Sales and marketing have heavily relied on creating relationships, and making guesses. Predictive analysis connected with big data is changing this entirely as it “*can finally become more about math than magic.*” (Hollison, 2015 p. 1). Insights gained from analyzing big data will benefit marketing professionals as they are now able to obtain a more precise target group. Furthermore, sales forecasting gains more precision continuously as well as the ability to monitor real-time sales through big data visualization. Startups such as Uber are able to greatly benefit from this as they include big data strategies in their business model from the start (Hollison, 2015).

Understanding user trends creates direct competitive advantage. The retail-firm Target has even gone so far that they can figure out which of their customers are pregnant through data mining (Watson, 2014). By assigning every customer to a unique Guest ID number connected to the customer's credit card, Target is able to identify user trends and create value. The mapping of the customers involves a complete history of everything they bought as well as demographic information. This not only creates a pattern, it maps out a complete picture of who the customer is.

2.5.2 Big data management

The organization, administration and governance of large quantities of unstructured and structured data is called big data management (Rouse, 2013). The goal is to achieve data of high quality that is accessible for business intelligence as well as big data analytics applications. Organizations, corporations and government agencies use big data management strategies in order to reach the goals. Since big data usually involves enormous sets of data, the strategy helps to structure and create value from the data. If the big data management is effective, the company can locate valuable information from a variety of sources including system logs, social media sites and call detail records.

Big data environments most often go far beyond the relational databases and traditional data warehouse platforms and are therefore more complex and more difficult to incorporate. Over time, more and more big data platforms are evolving and they are helping the user to understand which data is relevant and which is not. Ultimately it is used to improve current business processes and to create competitive advantage (Rouse, 2013).

By itself, stored data does not generate any value for the business. However if the data is stored appropriately, it can be used for analysis which can create great value (Watson, 2014).

2.5.3 Big data applications

We mention that big data can create great value if it is stored appropriately and used for analysis. Here we present three business examples to show how companies have been able to create value from big data.

Starbucks

As Starbucks was introducing a new type of coffee, a growing concern that the customers would find the coffee too strong arose. As the new product was beginning to sell across their different cafés, Starbucks started monitoring blogs, Twitter and various coffee forums in order to hear their customers' opinions. By mid-morning it was discovered that customers enjoyed the taste but thought that the new product was too expensive, a quick reaction from Starbucks led to a lowering of the price and the negative comments disappeared (Watson, 2014). This use of big data illustrates how analytics can lead to happier customers as well as more profit. A more traditional approach would have been to wait for the sales report for several weeks and then run the results by a focus group in order to find out why it sold less than expected.

Chevron

A drilling miss for Chevron in the Gulf of Mexico costs up to \$100 million. In order to improve the chances of finding oil, big data analytics are being used. However in 2010 the odds for a successful drilling was still only 1 in 5 even though they were analyzing 50 terabytes of seismic data. In the summer of 2010 as a result of the BP oil spill, Chevron put more focus on advanced computing power as well as data storage. After the data enhancements, the odds for finding oil increased dramatically. Now with the new technology, Chevron have been able to find oil 1 in 3 times instead of 1 in 5 which have led to great cost saving and more profit (Watson, 2014).

U.S. Xpress

U.S. Xpress is a transportation company with cabs constantly streaming out data regarding truck condition and location. This data is stored in the cloud and later analyzed in different regards. The information gathered is then delivered to various interest parties, from drivers to senior executives in a tablet format. If a truck is low on fuel, the driver is notified with information of the cheapest gas station and best located compared to the route he is driving. If it is apparent that the truck is in need of maintenance, the drivers are sent to specific service depots. By analyzing the routes and eventual idle time, U.S. Xpress are able to save millions in fuel costs as well as reducing the emissions into the environment by reducing idle time (Watson, 2014).

2.6 Big data benefits and opportunities

Big data is creating a new generation of decision support and businesses worldwide are recognizing the potential value gains. In order to succeed with big data, businesses are aligning technology, people and processes. Storing and collecting data does not create much value in itself. It is rather by analyzing and using the results found where decision makers and organizational processes are influenced. This is where the real difference is made and thereby value is created from big data (Watson, 2014). Moreover, big data is creating a high demand for work expertise in big data.

As mentioned earlier, collecting and storing data does not create any business value. Value is gained from analyzing and acting on big data results. The benefits from big data can be varied, substantial and the basis for competitive advantage. But the competitive advantage can only be realized when an organization has carefully thought out and executed a big data strategy (Watson, 2014).

Previous research by Management review in collaboration with IBM from 2010 show that the top performing organizations use analytics five times more than lower performers and that 37% of executives believe that analytics creates competitive advantage. One year later a follow-up study was made and the same belief had now risen to 58% (Watson, 2014). According to Watson (2014), this shows the positive relationship between data-driven decision making, organizational performance and competitive position.

Companies across different fields use big data to analyze and find trends, whether it is on a local or a global scale it helps to gain a better understanding of various fields of study (Cole et al., 2015). In the astronomy field, the Sloan Digital Sky Survey (SDSS) has provided and shown how big data together with computational methods can support and facilitate decision making at macroscopic and microscopic levels in a globalized and rapidly growing research area. As a result of big data, the SDSS is one the largest, influential and ambitious surveys in the history of astronomy (Chen et al., 2012).

Manyika et al. (2011, p. 4) claim that *“Big data (...) is now part of every sector and function of the global economy. Like other essential factors of production such as hard assets and human capital, it is increasingly the case that much of modern economic activity, innovation, and growth simply could not take place without data”*. Gantz and Reinsel (2011, p. 2) does not stop at companies, they also predict the change in society through Big Data: *“The convergence of technologies now makes it possible not only to transform the way business is conducted and managed but also to alter the way we work and live”*.

2.7 Big data risks and challenges

As proposed earlier from the McKinsey study, there will be a huge shortage of people with deep analytical skills and managers with relevant digital competence (Chen, Chiang and Storey 2012). Big data opens up many opportunities for companies but collecting such vast amounts of data is not risk free. What happens if unauthorized individuals gain access to the data? Are the huge corporations being transparent enough

about how the collected data is used? Due to the fact that big data still is a relatively new phenomenon there are still many uncertainties.

Cloud Computing

A big amount of the risks in correlation with big data is when cloud computing is involved. Outsourcing to cloud service providers is widely used since it lets the company focus on its core business (Cole, Nelson, McDaniel, 2015). When data storage is relied upon others, obviously the control over the data is decreased or even lost completely. *“Confidentiality, system integrity, data integrity, and reliability may all become a greater risk when total control over data is lost”* (Neumann, 2014, p.1).

Data Ownership

Concerns with big data arises when there is uncertainty concerning who owns the data and who has rights to the data (Cole et al., 2015). Companies want to believe that they control and own the data that they collect. This is often not the case. Individuals on the other hand are interested in how much information about them is stored in the various databases. For the individual, big data poses a threat to privacy as well as a data rights issue. Furthermore the individuals want to make sure that whatever is stored about them is accurate and correct (Cole et al., 2015).

Personal Integrity

In most cases the individual has to agree to have data collected from them. Although it is more rare that the individual actually signs a sheet of paper. One of the biggest collectors of data, Facebook, also requires the user to agree before they collect the data. This is agreed upon when the user signs up for an account but most people do not pay attention to this detail. A study conducted by Bechmann (2014) suggests that the informal consent button creates an imbalance between user and service provider in the favor of the service provider. The study also suggests that individuals using these services cannot comprehend the extent to which the service provider, Facebook, can use data including privacy inbox and news feeds (Bechmann, 2014). Using the button does not feel the same as physically signing personal information over to another; therefore, the user may not feel the need to read the fine print.

Data Quality

When it comes to security, the cases are mostly related to data hacking and sensitive data exposure. Security can be split into two parts, internal and external. Internal security covers employees and in-house company issues. External security is about the system itself, the structural risks and possible vulnerabilities. A big issue is the misinterpretation of big data. While the large amounts and difficult structuring of the data mitigates other risks concerning security, it increases the risk of misinterpretation. Misinterpretation of trend analysis, sales projection and user behavior can lead to catastrophic results. Therefore it is important to use a stable and secure system. Many low budget open source tools are therefore advised against when working with big data (Westermeir, 2005). Large companies have an automatic advantage in this sense since they have a larger budget and can therefore purchase a stable and bug-free system. An old issue that is still the biggest threat to security is the human factor considering big data. Internal security is still to this day the greatest risk (Spears and Barki, 2010).

2.7.1 Issue areas

To further elaborate the risks and challenges regarding big data, Gao et al. (2015) address three different issue areas:

People

Experts within big data are rare and therefore expensive. Simply enough, there are too few with this kind of expertise within computers and analytics. These experts are referred to as data scientists (Gao et al., 2015). In big data projects, it is vital to include these experts from the start in order to gain maximal value (Sicular, 2012). Furthermore, it will be more beneficial to hire more than just one data scientist in a big data project since they require a whole set of expertise skills from different fields. The strongest results are achieved by having a multidisciplinary analysis team with the skills required for the specific project (Sicular, 2012). Sicular (2012) further claims that risk management and legal experts should be included early in the project since there are strong ethical, security and legal issues regarding big data.

Process

Big data projects are innovation projects and therefore require the project teams to be able to quickly react and adapt to sudden changes. The project team must be set up in a way so that they are flexible and ready for changes. Revealed data management problems are discovered through big data analysis and require immediate action (Gao et al., 2015).

Technology

Since the big data era is relatively young there is an absence of end-to-end solutions, design patterns and universal architecture. As a result from this, big data projects and the technology are forced to be individually assembled and customized to the purpose of the analysis project (Sicular, 2012). Moreover, because of the enormous amounts of data used in the projects, storing is eventually going to become an issue and this needs to be addressed early in the project.

2.8 Requirements for success with big data

To further understand how big data can be used and create value in the corporate context, we present an overview of important factors.

2.8.1 Big data strategy

Gao et al. (2015) claims that a big data strategy should not replace existing traditional analysis methods, instead it should work as a complementary and extend the possibilities of the existing methods. This is because the traditional analysis of structured data still creates value for the business. The big data strategy has to ensure that the project is aligned with the business. Therefore, this means an alignment of the people involved in the project and the technology used in order to support the business goals (Gao et al. 2015).

A big data strategy is also used in order to help organizations foresee where valuable insights can be found. The strategy is further used to avoid the early big data challenges for instance by avoiding the issues regarding storage of data. The second part of the big data strategy is the integration with the organizational strategy. An understanding of the business strategy and to choose the correct analytical toolset thereafter is one of the major keys to a successful big data project. An alignment of management and organizational culture to be able to leverage from new insights gained will help organizations to utilize the full potential of big data (Gao et al., 2015).

Gao et al. (2015) also focuses and illustrates the importance of asking the right questions. This means strong analytical talent and the involvement of “the right people” in order for a big data project to be truly successful. There is no point gathering a big amount of data if the right questions are not being asked (Gao et al., 2015).

2.8.2 Key terms for success

In order to succeed with big data analytics, Watson (2014) proposes several key terms relatable to our research question. However, he also mentions that many of the key terms for success with big data projects are largely the same as with most other projects including analytics and Business Intelligence. The differences are found in the details and some of the details are very important (Watson, 2014). We have chosen to highlight a couple of these key terms that we find relevant to our study.

A clear business need

Projects should be business driven rather than technology driven. Projects should address a certain business need, like solving a problem or seizing an opportunity. Watson (2014) means that because of the media hype of big data, executives have pushed big data projects without any clearly defined business goals.

Alignment between the business and analytics strategy

Big data projects must support the business strategy, therefore most projects should be driven by business people rather than by IT. In some organizations, the alignment is even more important. In analytics-based organizations it may even be impossible to separate the analytics strategy from the business strategy. Without analytics as the enabler the business strategy cannot succeed (Watson, 2014). A visible example of analytics doing work are product recommendations which differ between customers on web sites. The recommendations are a result from customers' previous mouse clicks, other purchases, availability of products, profitability of products as well as search terms entered in the past.

A fact-based decision-making culture

For an organization to benefit from big data, decisions must be based on facts according to Watson (2014). These facts are generated from big data analytics and the decision-makers must be able to trust this data in order to make the correct decision. To change an organizational culture can be very challenging and even more challenging than solving technical issues. Creating a fact-based decision-making culture might involve replacing people who cannot or will not adjust to the new culture (Watson, 2014).

People skilled in the use of analytics

It is important to have people with the relevant and necessary skills to work efficiently with big data. Gaining value from big data requires a big process and many different experts including business, data and analytics expertise. The business users access the big data information through reports, data visualization and different kinds of dashboards or scorecards. Watson (2014, p. 17) means that "*They are information consumers rather than information creators*". The business users do not need full understanding of all technology, instead they need to understand the possibility of big data analytics. They need to understand what data is available, how to access it, how to manipulate it in simple ways and how to use it to create business value (Watson, 2014). A data scientist's role however is different, they need the expertise in order to find trends, patterns and relationships and then provide the data to the business user and thereby create value for the organization.

2.9 Big data sources

Big data is not just a trend, it is an ongoing development and the amount of data is expected to grow even faster in the future (Gao et al., 2015). One of the technologies contributing to the rapid growth of data is the internet of things. The network of connected physical objects embedded with sensors, electronics and software gives organizations access to massive chunks of real-time data. This relatively new technology is opening up endless opportunities for a range of innovative applications (Chen et al., 2012). The devices are enabling flexible and mobile usability as well as a new way of data collection and insights. These

highly mobile, location-aware, person-centered and context-relevant transactions and operations will continue to create unique opportunities for researchers in the future (Chen et al., 2012).

Sensors integrated in smart phones, cars, credit cards and general consumer electronics are providing a completely new range of data to be gathered. Furthermore, social media is streaming out endless amounts of valuable customer information to companies. Because of this, companies are more or less forced to address these new data sources if they want to stay competitive in the big data environment (Gao et al., 2015).

Every single mouse click on web sites by every single user can be captured in web log files and can later be used for analytics. Typically these clicks are being analyzed by companies in order to better understand online shopper's behavior and create better user experiences thereafter with dynamically changing product marketing (Watson, 2014). Social media such as Twitter and Facebook are generating enormous amounts of comments and tweets which as well are being analyzed to better understand customer trends and to improve products.

Machines such as smart meters are also generating streamed data about electricity levels, water levels or gas consumption etc. which can be shared with the customers or utilized to improve pricing plans. Another source is geospatial data such as GPS data created by smart phones (Watson, 2014). Uber is a clear example of a company successful of using this kind of data. Other examples are coupon based apps that give their users special deals depending on their location.

2.10 Other use-areas of big data

There are many different applications of big data. As we are putting our focus on IT capabilities and value creation we still feel that we need to include a couple of other use-areas of big data.

Smart Cities

Another upcoming trend is to make cities smarter. As part of creating a smart city, municipalities use various tools to improve the city's infrastructure, CO2 emissions, the overall living standard and more. One of the tools used to achieve this is the internet of things. By putting up sensors around the city in everyday objects like lampposts, the municipality monitor the traffic situation, measure the air pollution or automatically adjust the lights after the outdoor brightness. The sensors could also be put into trash bins that sends an alarm when it is almost full in order to make the trash pick-up route more efficient. There are also sensors that can monitor the water-level in cities with flooding problems. This feed of data can be used as a countermeasure to reduce the overall damage from the floodings. The same thing can be done with cities affected by earthquakes where a more accurate warning can be received in an earlier stage. All this data paves the way for how our future cities will look like. However it puts further demand on the skill to properly analyze and interpret data.

Healthcare

Big data is creating enormous opportunities across a vast amount of fields as mentioned earlier, one of the most important fields is the health community. Just like every other field they stand before supreme amounts of collectible data which potentially can lead to revolutionary insights. Two main sources of medical big data are discussed by Chen et al. (2012). Firstly there is genomics-driven big data such as genotyping, gene expression and sequencing data. Secondly there is payer-provided big data such as electronic health records, insurance records, pharmacy prescription and patient feedback and responses. Electronic Health Records (EHR) have been implemented across various hospitals around the world during the last 15 years which in turn are helping analysts gain better knowledge about diseases and find potential trends. Furthermore in addition to EHR, several health social media sites have emerged. These

social media sites provide great research opportunities and better decision making with accurate social media analytics (Chen et al., 2012).

In the medical research field, medical professionals and data scientists are now able to study more patients at a higher speed and accuracy which increases the research efforts. For instance, big data is analyzed to map global infectious disease and is working as an offensive strategy to battle the disease. This strategy makes it possible to find the important trends which can lead to an understanding of the geographical range of the disease. Thereby it is now easier to estimate outbreaks and prevent them (Cole et al., 2015).

Big data and social media

Business analytics has become a top strategic priority for senior executives. Business analytics systems enables managers and other decision makers to analyze and interpret data to improve decision-making as well as business processes (Kiron, Ferguson and Prentice, 2013). As social media is creating huge amounts of unstructured and semi-structured data, managers are starting to realize the potential value within this data. Social media has revolutionized the way organizations are communicating with their customers and environment as it has created a new channel of communication (Kohavi, Rothleder and Simoudis 2002). Consumer opinions, experiences and sentiments towards brands, products and services are potential insights found from analyzing big data within social media.

A clear example of a company utilizing big data is Facebook. Many probably know that just like Google, Facebook makes most of their profit from advertising. What differs a Facebook advertisement from a regular television commercial is the target scope. Facebook enables the selection of complete demography. This means that the advertiser can choose exactly who they are targeting. City/country, age, sex, interests, relationship status as well as education and work are factors the target group can be chosen from (Facebook, 2016).

2.11 The Resource Based View

The Resource Based View (RBV) is a theoretical perspective that convincingly addresses the complementarity of IT and organizational processes, practices, routines, and activities (Wernerfelt 1984, Barney, 1991). RBV argues that durable competitive advantage emerges from unique combinations of resources (Grant, 1996) that are economically valuable, scarce, and difficult to imitate (Barney, 1991). From this perspective, there are compelling theoretical reasons for investigating how firms allocate investments across different types of IT assets (Aral and Weill, 2007).

Wade and Hulland (2004) argues that the RBV indeed is useful to IS research. It provides a valuable way for researchers to acknowledge how information systems relate to firm strategy and performance. *“The theory provides a cogent framework to evaluate the strategic value of information systems resources. It also provides guidance on how to differentiate among various types of information systems - including the important distinction between information technology and information systems and how to study their separate influences on performance”* (Wade and Hulland, 2004, p. 4). RBV also provides the opportunity to compare IS resources and non-IS resources. However it is important to note that IS resources do not contribute sustained competitive advantage (SCA) on their own. They can instead be seen as a part of the “big picture” of a firm which lead to competitive advantage, necessary but not sufficient (Wade and Hulland, 2004).

The resource based view has been widely used and interpreted among information systems researchers. It has been used to understand the link between investments and organizational value as well as competitive advantage (Bekmamedova and Shanks, 2014) (Kohli and Grover, 2008). From the resource based view, an organization can be looked upon as a collection of resources that enables it to compete and succeed. Resources in this sense are assets and capabilities merged together. The RBV has been used in several

previous studies in order to understand the impact of various IT investments across different levels in the organization (Bekmamedova and Shanks, 2014).

2.11.1 Resources

“Resources are the foundation for strategy, and unique bundles of resources generate competitive advantages that lead to wealth creation” (Hitt, Ireland and Hoskisson, 2010, p. 72). Resources can be seen as a collection of individual, social and organizational phenomena. Resources are divided into tangible and intangible resources. Tangible resources are traditionally assets that are quantifiable and visible.

Hitt et al., (2010) divides the resources into four different categories: financial resources, organizational resources, physical resources and technological resources. In contrary to tangible resources, intangible resources are assets rooted deeply in the firm’s history and have developed over time. Since they are hard to define, built upon several years and embedded within several years of routines and patterns these resources are hard to imitate and analyze. According to Hitt et al., (2010) there are three types of intangible resources: Human resources, Innovation resources and reputational resources.

By measuring tangible resources one can find the economic status of an organization but it does not provide a complete picture of all valuable assets since it disregards the intangible. Furthermore, production assets are surely important but they create no value without the relevant processes (which are intangible). This shows the importance of the correlation between tangible and intangible resources (Hitt et al., 2010).

For resources to be of any strategic importance and achieve competitive advantage, they must be valuable, rare, inimitable and non-substitutable (VRIN) (Larson and Watson 2011).

2.11.2 Assets

We have chosen to use the definition published in 2006 by the International Accounting Standards Board (IASB). An asset is a resource controlled by an entity as a result of past endeavors and from which future economic benefits are expected to be derived. Even though employees can produce future economic benefits they are not considered to be assets, due to the fact that human capital cannot be fully controlled by the entity (IASB, 2006). There are both tangible and intangible assets. Tangible assets refers to those that have a physical substance, this could include buildings, company cars and company equipment. The intangible assets on the other hand refers to non-physical objects like copyrights and branding.

2.11.3 Capabilities

Purposely integrated resources that have been implemented in order to achieve a specific set of tasks are called capabilities (Hitt et al., 2010). The capabilities can range from product marketing, human resources and research & development activities. Hitt et al. (2010) states that for organizations, the competitive advantage is often based on their human capital and their ability to exchange knowledge and information. Furthermore they state that the employees’ unique skills and functional expertise lay the foundation for many capabilities. While the knowledge withheld by the human capital poses a significant value for businesses, it is at least as important to be able to utilize this expertise. For organizations this poses a challenge of creating an environment that facilitates knowledge transfer between employees and the different business units (Hitt et al., 2010).

There are four core capabilities required for a company to sustain competitive advantage (Hitt et al., 2010). The first one is **valuable capabilities** which allows the organization to neutralize threats and exploit opportunities in its external environment. Another core capability is to have **rare capabilities** that only a few, or none, of the competitors possess. **Costly to imitate** is another core capability that refers to capabilities that are hard for competitors to replicate. This might include capabilities derived from unique

historical conditions, the corporate culture or the link between a firm’s capabilities and its competitive advantage. The last core capabilities are the **non-substitutable** ones which have no strategic equivalent, this could for instance include culture and brand-image (Hitt et al., 2010).

2.11.4 IT Resources

IT resources are a combination of IT assets and IT capabilities. Merging together the allocation of investments with strategic purposes and a system that mutually reinforces practices and competencies (Aral and Weill, 2007). The relationship between IT investments and different measures of firm performance can be moderated by organizational capabilities. However if an organization has obscure definitions of their IT resources and aggregated measures of their IT investments it might lead to various results (Aral and Weill, 2007). For instance organizations with a greater e-commerce capability have a better performance in areas like supply chain optimization but perform worse when it comes to the cost of goods sold (Zhu and Kraemer 2002).

The landscape of IT resources in firms varies, depending on how they allocate their investments. Making heterogeneous investments in the pursuit of different goals, like innovation or cost leadership is one factor that could separate one firm’s IT resource landscape from another. A company who invests their assets in the pursuit of innovation strategies is likely to invest more into systems that support process and product innovation, while companies with a cost leadership strategy is probably focusing more on investing towards transactional IT systems. Another factor that separates firms’ landscapes from another is their IT culture (Aral and Weill, 2007). Aral and Weill (2007) also states that it is difficult to attain strong IT resources due to the interplay between IT assets and IT capabilities. Achieving strong IT resources takes time learning and optimizing.

The theoretical model presented below was constructed by Aral and Weill (2007) with an aim to “sharpen the theoretical characterization of IT resources by unpacking two major sources of variation in the empirical evidence on complementarities between IT and organization: heterogeneity in IT investment allocations and organizational IT capabilities.” (Aral and Weill, 2007).

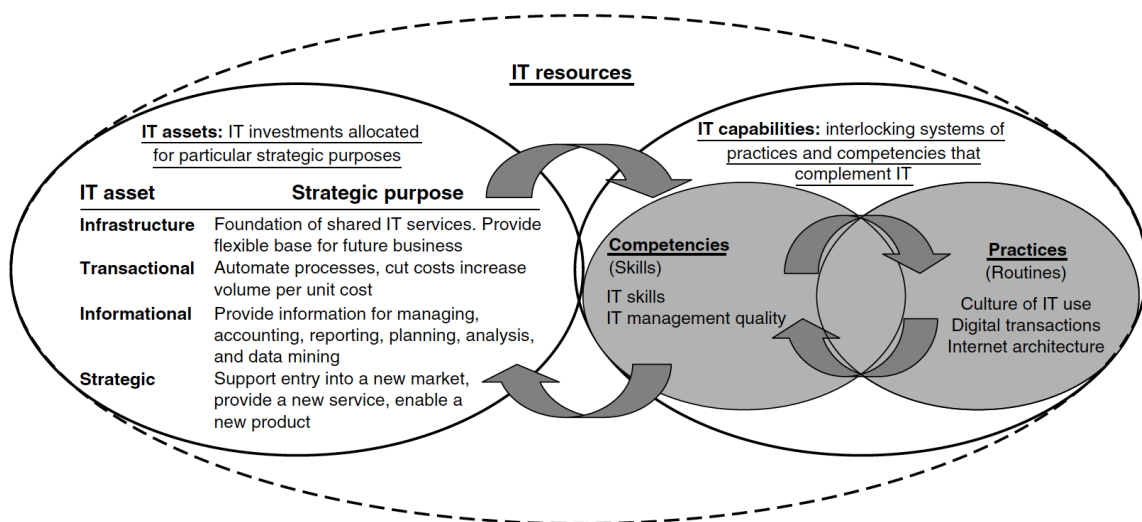


Figure 3, Theoretical Model of IT Resources (Aral and Weill, 2007)

We decided to build our theoretical model based on the model presented in figure 3 since we feel that it provides a good overview of what IT capabilities are in relation to IT resources and IT assets. We also define big data as an IT resource, a combination of IT assets and IT capabilities.

2.11.5 IT Assets

IT Assets in this sense are based on a framework developed by Weill (1992). The framework was later extended by Weill and Broadbent (1998). The framework situates IT investments into four different categories of assets: Infrastructure, transactional, informational and strategic assets.

1. *IT infrastructure* provides the foundation of shared IT services (servers, networks, desktops, laptops etc.) which are used by several IT applications (Keen 1991). Infrastructure investments are typically made to provide a flexible base for future business initiatives and are thus made in anticipation of future business needs. Infrastructure investments enables new applications and functionality and can be seen as building blocks in order to create long term improvements (Duncan 1995).

2. *Transactional investments* are used to reduce costs, automate processes and increase the volume of business a firm can conduct per unit cost. Order processing, cash withdrawal and other repetitive transaction processing functions are examples.

3. *Informational investments* are used to provide information for accounting, managing, reporting and communicating internally as well as with customers, suppliers and regulators. Sales analysis, decision support and planning systems are examples of informational investments. By analyzing sales, customer trends and product reaction it helps managers take better decisions. Aral and Weill (2007) expect informational investments to tighten reporting, control functions as well as improve data collection and decision making, thereby reducing costs and identifying new opportunities for revenue generation and profitability improvements.

4. *Strategic investments* are used to reposition firms by enabling or supporting entry into a new market or the development of new products, business processes and services. A successful strategic investment changes the nature of the product or service. However, it becomes non-strategic if and when competitors are available to commoditize the capability (Aral and Weill, 2007). An example of this is the ATM in the banking industry. The ATM completely changed the nature of service delivery and created competitive advantage for the early adopters (Dos Santos and Peffer, 1995). However when more and more competitors started adopting it, it became non-strategic and transactional.

2.11.6 IT Capabilities

Aral and Weill (2007) claims that a variety of individual capabilities, practices, and processes may complement IT; however, they expect systems of practices and competencies working in concert to enable greater business value. By using prior research and creating case studies they identified a group of interlocking organizational characteristics that together support firms' ability to derive value from IT (Aral and Weill, 2007).

From Nelson and Winter's (1982) original theories they defined Competencies and Practices as two fundamental conceptual building blocks useful for identifying characteristics of firms that complement IT (Aral and Weill, 2007).

Competencies in this sense are skills embodied in employees or groups that actively manage or accomplish organizational tasks (Pralhad and Hamel, 1990, Dosi et al., 2000). Competencies are developed through the repeated performance of contextual activities and learning. As individuals and groups interact with IT for particular purposes, they learn, build skills, and develop competence toward effective use (Aral and Weill, 2007).

Practices, in contrast, refer to recurring sets of activities or routines that serve both as a means of accomplishing organizational tasks and as mechanisms for socially storing and accessing knowledge about the most effective ways to accomplish those tasks (Cohen and Levinthal 1990).

Practices and competencies support and complement each other. Practices help individuals and groups develop competence or skill with particular ways of working (Dosi et al. 2000), and skills are necessary for the effective execution of organizational practices toward specified goals (Nelson and Winter 1982).

Further, competencies in this case are defined to exist across two organizational dimensions. The IT skills of employees at all levels (both technical and business skills) and IT management competence (Aral and Weill, 2007). IT, big data and business understanding skills are related to application development, database administration and networking in relation to business skills (Bekmamedova and Shanks 2014). In the case of data analytics, skills required include management of unstructured data, text mining, natural language processing, machine learning and social network analysis (Chen et al., 2012). Moreover Santhanam & Hartono (2003, p. 4) states “*firms that develop an effective IT capability are able to obtain superior financial performance compared to those who do not develop an effective IT capability.*”

An alignment between different areas of the organization such as sales, marketing, product development and IT are important to gain insights and create value from collected data. Organizations which have digital work practices ingrained in their employees work routines are more likely to succeed when it comes to using big data solutions (Aral and Weill, 2007).

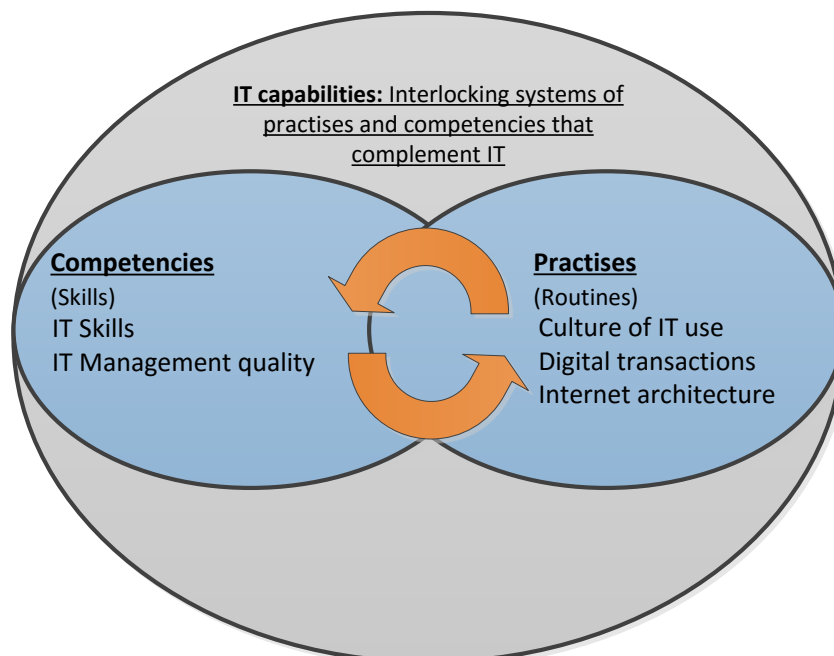


Figure 4, Theoretical Model used for our research. Based on Theoretical Model of IT Resources (Aral and Weill, 2007).

Presented in figure 4 is the final theoretical model we have chosen to base our research on. The model is based on the IT resources model created by Aral and Weill (2007). The model depicts the important correlation between competencies and practices which together makes up the IT capabilities of a company. The model is then applied to get a better overview of what IT capabilities are required to create value from big data. Since we are mainly focusing on IT capabilities and value creation we have chosen to exclude IT assets from the model. However we still want to highlight that IT capabilities on its own is not

enough to show why a company can create value from big data. Therefore we try to remind of the importance of IT assets in correlation with IT capabilities which together create IT resources.

2.12 Literature Summary

We have presented the most relevant literature we have used and proposed our definition of big data. Further we present a model showing the definition we have used for the big data process. We also say that we focus on big data in the corporate context and motivate why. We further present how big data is changing sales and marketing, what big data management is and give three business examples of how big data has helped three different companies create value: Starbucks, Chevron and U.S. Xpress.

To gain a better understanding of how big data is affecting business, we present the benefits and opportunities as well as challenges and risks of big data. We then introduce different big data sources so the reader can understand the immensity of big data. As our thesis treats value creation of big data, we have included requirements for success with big data. We further propose the big data strategy and key terms for success with big data.

As we start to explain why we have chosen our theoretical model, we start by introducing the RBV. The RBV works as a ground pillar showing the reader the perspective we apply looking at companies and their work with big data. Additionally, we give brief definitions of resources, assets and capabilities. This is done in order for the reader to gain a better understanding of IT resources, IT assets as well as our main focus; IT capabilities. We also say that we see big data as an IT resource. The finalized theoretical model we apply in our research is built upon the RBV and the model in figure 3 created by Aral and Weill (2007). Furthermore we show how the model is connected to the value creation of big data.

3. Method

In this chapter we will describe and motivate the method that we have used for our preparatory work, data collection and analysis of our empirical material. We have used Jacobsen's book 'Vad, hur och varför? - Om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen' to decide how to approach the data collection of the thesis.

3.1 Data collection method

The goal of an empirical research is to investigate and depict the reality (Jacobsen 2002). To form a theoretical basis and make an empirical study of it, Jacobsen (2002) talks about the importance of interviewing the right people. People who have experience and have a good insight regarding the subject. In order to give a fair judgement and contribute with research of value we deem it of high importance to get in touch with the right people. To get in touch with the right people we have contacted companies through phone and email and tried to reach people with different roles in the firms' big data process.

According to Jacobsen (2002) the two primary methods used while collecting data for a research papers is either quantitative or qualitative. We chose to use the qualitative method due to the fact that through interviews we can get achieve a deeper understanding of the capabilities required for extracting value from big data instead of just a general overview. In order to achieve a deeper understanding, it is important for us to be able to adapt our questions to the answers they provide instead of sending out a pre-set questionnaire.

Jacobsen (2002) describes that it is important to have some kind of structure while designing the interview guide. Having a structure facilitates the analysis of the collected data. We tried to form the interview guide in a manner that allows us to have an open conversation where the interviewees can voice their own opinions. Through the interviews we wanted to create an empirical foundation from which we could easily compare and analyze the results.

3.2 Company and participant selection

The selection of companies is based on industries that actively use big data as part of their daily operations. We have focused on companies with customer-oriented operations. Customer-oriented companies have a great interest in collecting data about their current and potential customers. The companies we have chosen either benefit directly from big data or have the opportunities to benefit from big data solutions. All companies chosen have a strong connection to IT, data collection and analysis.

Softhouse

Softhouse is a consulting firm that was founded in 1996 and is today one of the leading companies in the IT industry in Scandinavia. Softhouse has about 200 employees and their area of expertise lies in Lean and Agile methodologies. The company has an impressive history of constant growth and has also won several awards. The Travel Planner in Skånetrafiken is one of many products that Softhouse has developed. At Softhouse every customer has a salesperson who reports the data of customer relations, customer visits etc. This data makes it possible for decision-makers to see who is on their budget and who is not. Every month, business reports are made from the collected data by IT personnel. These reports are then shown to the business people as an overview from which they can get important insights and take better decisions.

Avensia

Avensia is a leading company within the field of e-commerce. They deliver complete so called Omni-Channel (e-commerce) solutions to companies with high demands. Sales statistics, order history, customer behavior, customer statistics, segment statistics and logistics flow. These and more are what Avensia

collects for their customers. Afterwards, they form their e-commerce websites from these different results and trends found. Of all the millions of customers that visit the different sites, they track every single one. Consumers' mouse clicks leading them to the website, what they have shown interest for, what they put in the shopping cart, what they bought and what they looked at are other examples of data that is collected. This is then used to optimize every individual customer's web site experience by comparison with other former customers. The customer experience is continuously improved by adjustments in real time. This kind of collection is automated, however the process of binding together and then analyzing the results is semi-automatic. Lastly when the analytics are done, targeted campaigns and other similar actions can be made from the collected insights and trends found.

King

King is a marketing firm known for working with some of the strongest brands of Sweden. They are mostly known for their impressive, creative and hilarious TV-commercials. King gets hired from customers who want better communication. What King does is to tell their customer what their target market looks like, which competitors they have and how those work and act. Further they explain who their consumers are, how they work and act, this is their perception of the brand, these are the current trends and so on. *"If a customer wants to talk about meat we will do an analysis to figure out how the competitors are talking about it. Then we might find a possibility to talk about it in a different way, or perhaps all the customers have recently watched Cowspiracy and do not want to talk about meat at all. Then perhaps you should address it a completely different way, who knows..."* (Line 25, Appendix 6).

Lundalogik

Lundalogik is one of the top CRM-providers in Scandinavia. They have three products, all within the area of sales and CRM. Internally they are collecting data from all the Nordic companies and their contact persons. They have a system that keeps track and updates the contact details of all the companies' decision makers every day. These big chunks of data is being used in order to find new potential customers.

Styleport

Styleport is a freshly started e-commerce company that has yet to present its first annual report. With only a few employees their work with big data is still modest. As far as data goes they are currently only working with their customers phone numbers and mail addresses to send newsletters. However they plan to implement big data solutions in the near future which is why we considered it interesting to include them in our research.

3.2.1 Participants

As we wanted a holistic perspective of the big data process we looked for interviewees with varying roles connected to the companies' work with big data. Our target roles have been: CEOs, CIOs, head of marketing and data analysts. Since all the companies have a strong connection to IT, all the interviewees have been IT savvy. Due to the relatively small size of the companies some of the interviewees has held several titles. The CEOs interviewed serves a multi-functional purpose where they also act as CIO or as a head of marketing. The planner was interviewed due to the fact that he also had the same role as a data analyst of his company. We chose to interview data analysts in order to get a better understanding of how the data is used and how the communication between the analysts and the decision maker works.

| Company | Line of Business | Employees | Revenue(SEK) | Position |
|------------|------------------|-----------|--------------|--------------|
| Softhouse | IT-consulting | 173 | 186 millions | Data analyst |
| Avensia | E-commerce | 73 | 112 millions | CEO |
| King | Advertising | 30 | 130 millions | Planner |
| Lundalogik | CRM-systems | 165 | 127 millions | Deputy CEO |
| Styleport | E-commerce | N/A | N/A | CEO |

Table 2, Companies interviewed

3.3 Conducting the interviews

The interviews have been conducted through phone. The fact that the phone interviews are less personal does not have to be a bad thing. According to Jacobsen (2002) a phone interview can reduce the interviewer effect (Groves & Kahn 1979). The interviewer effect implies that the physical presence of the interviewer can cause the interviewee to act abnormal. The interviewee can observe the interviewer's facial expression, body language etc. If the interviewee for example notices that the interviewer reacts in a certain way after an answer or is bored, the interviewee can adjust their answers accordingly and cause the interview to become less open (Jacobsen 2002).

All the conducted interviews have been recorded digitally. This has been done in order to ensure that no information is lost. Having all the interviews transcribed word for word will facilitate the analysis of our data.

3.3.1 Interview guide

The interview guide we created has been designed to facilitate the analysis of the collected data and to allow us to gain a deeper insight. It is important to have some kind of pre-structure of the interview in order to cover all the aspects that we want to research (Jacobsen, 2002). Jacobsen also states that without any kind of structure of the interview it is easy to miss important subjects or not get the most out of the answers from the interviewee. However having a too strict structure could result in loss of valuable data if the interviewee is not allowed to speak freely (Jacobsen, 2002).

Following Jacobsen's advice we created a semi-structured interview guide with space for open talk and follow-up questions. We have started with easy questions like letting the interviewee explain their work role. As the interviewee became more comfortable we started asking the bigger and more complex questions. According to Jacobsen (2002), starting with complex questions might result in the person giving less informative answers and it might also affect the flow of the remainder of the interview.

We based our interview questions of the theoretical model presented in figure 4. During the interviews our main focus has been to investigate what IT capabilities the companies have and how they are utilized in the work with big data. We have tried to investigate if the companies have the resources to handle their data efficiently. Furthermore we want to get a picture of what value potential the companies see in big data.

In regards to our interview questions we have chosen to focus on two main areas: IT capabilities and value creation of big data. We also chose to include questions regarding company background and the interviewees' background. The purpose of our background questions is to achieve an understanding of

how the company works with big data and why they chose to start working with it. We also included a few questions about IT assets since it is still a vital part of the value creation of big data. The purpose of those questions is to give us a better understanding of the correlation between IT assets and IT capabilities.

Most of our questions have covered our two main topics. The questions were also designed based on the research presented in our literature review. Since we used a semi-structured interview guide, the questions differed slightly between the different companies as we asked several follow-up questions based on the interviewees' answers. Our goal has been to formulate questions in relation to answering our research question: *What IT capabilities does a company need in order to create value from big data?* We see the big data process as an IT capability since we define it as an alignment of competencies and practices. We have investigated what IT competencies are required from different key roles in the big data process since this is one of our main definitions of an IT capability. Furthermore the questions are designed to investigate the opportunities big data has opened up for organizations.

3.4 Data Analysis

The process of analyzing the collected empirical data consists of the phases; *description, systematization and categorization* as well as *combination* (Jacobsen 2002).

Description: As a researcher it is important to get as thorough and detailed description of the data as possible, without spoiling it with our own input.

Systematization and categorization: The next phase is to systemize the information and to create an overall perspective in order to make it easier to understand.

Combination: When the data has been systemized it is time to interpret the data and try to find commonalities which we can derive conclusions from.

In a qualitative research the three phases are not sequential. It is possible to plan, conduct interviews, make observations and analyze them and then go back to the planning phase to adjust the setup (Jacobsen, 2002). According to Jacobsen (2002) this facilitates an iterative process of adjusting our method to the accumulated knowledge gained from the research process. We have chosen to use this approach as it makes us more agile. After every interview we revised our questions and adjusted them to our acquired knowledge; removing redundant questions and adding new ones that will help us better depict the reality. Although after the two first interviews there were only minor changes to our interview guide. In order to not miss out on any important data we made sure to contact the two first interviewees and complement with our additional questions.

The *description* of the collected data was done through transcription of the recorded interview. In order to get as much data as possible out of the interviewees we have been asking follow-up questions when we felt that we needed more data to get a correct understanding of the situation. During the interviews we have both been asking the questions to complement each other and to allow both to get a complete understanding of the data. After finishing the transcriptions we have emailed them to the interviewees in order to ensure that they are correct and that we do not expose any sensitive information.

In the phase of *Systematization and categorization* we structured and compared the collected data. We structured the data in different categories, Background, IT capabilities, Value creation and IT assets. These categories are based on our general questions in order to get a better overview. With the overview we could derive the commonalities and try to make out patterns.

In the analysis phase of *combination* we looked at the categories and tried to find commonalities and discrepancies in the data. When looking at the structured data we were able to find the more subtle information and draw conclusions from it. We also used quotes from our interviewees that would give the thesis more color and further strengthen our conclusion.

3.5 Research Quality

We have tried to ensure a high data quality throughout our research. In order to keep a high data quality it is important that the thesis and its conclusions becomes subjects to critical review (Jacobsen 2002). To ensure the quality of the thesis we have tried to take into account the validity and reliability of the study.

3.5.1 Validity

There are two different types of validity; internal and external (Jacobsen 2002). The internal validity focuses on the validity of the resource; if the description of the phenomenon is valid (Jacobsen 2002). It practically means that a description only is valid if multiple individuals agree on the description, the more that agrees on it, the likelier that it is valid (Jacobsen 2002). We have made sure that the conclusions that we have drawn have been cross-referenced with our interviewees and previous literature. By sending the transcriptions back to the interviewees we also give them the opportunity to point out any information that has been misinterpreted by us, further increasing the validity.

While the internal validity is about describing a phenomenon in a correct way; the external validity is more about determining to what extent the observations from the research can be generalized (Jacobsen 2002). The purpose of qualitative methods is not to draw a general conclusion from a select group and apply it to a broader spectrum of people. The purpose is rather to understand and deepen concepts and phenomena (Jacobsen 2002). Since our research only includes data from a few select entities; we are not able to perform any big generalizations. However we have found patterns and similarities from our data which might be subject to future quantitative studies.

3.5.2 Reliability

Another aspect of ensuring the quality of the research is to investigate how reliable the gathered data is. According to Jacobsen (2002) it is important to critically examine a few different effects in order to determine the reliability of the study.

The first effect to consider is the interviewer effect; which means that the interviewee is affected by the interviewer. The interviewee is influenced by clothes, body language and how the interviewer speaks. This could have an impact on how the interviewee chooses to answer the questions (Jacobsen 2002). Through having all interviews over phone we reduce the impact of the interviewer effect but we still had to be mindful of how we chose to phrase ourselves in order to not distract the interviewee.

The second effect to consider is the context effect; which according to Jacobsen (2002) describes the effect as to what kind of dimensions of the context that could impact the results. There are two kinds of context effects, artificial and natural. An artificial context means that the interview is conducted in a context that is unusual for the interviewee, while the natural context means that the interview is conducted in a familiar environment. We have tried to neglect the effects of the context effect as much as possible. By having all the interviews conducted through telephone, the interviewees have been able to choose their own environment. We have also been flexible upon deciding when to perform the interviews in order for the interviewee to be in a comfortable context. Before every interview we have sent the general questions to prepare the interview of what is coming.

Besides from these two effects there is a third final threat to the reliability of the thesis, lack of attention from the interviewers during the interviews (Jacobsen 2002). This means that the data provided by the interviewees has not been recorded completely and could thus be misleading the derived conclusions from the research. To avert this we have asked our interviewees of their permission to digitally record the interview and clearly informed them of what the data is going to be used for. By having the data recorded digitally we did not have to rely on our brain capacity and our written notes to ensure that no significant

data is lost. However we still had to stay alert in order to ask relevant questions and get as much as possible out of the interview.

3.5.3 Ethics

When performing a research there are several ethical aspects to take account for. As explained earlier, the reliability of the research could be threatened by people knowing that they are being observed. The interviewees could be acting and answering different from how they normally would in order to satisfy the interviewers or to hide their true self (Jacobsen 2002). Jacobsen (2002) states that there are three basic requirements which a research should satisfy: *informed consent*, *demand on confidentiality and privacy* and *demand of an accurate review*.

The basis of *informed consent* is that the interviewee is willingly participating in the research and that the participant knows about all the gains and risks that participation would entail (Jacobsen 2002). One of the factors that might impact the volunteering of the interviewee could be that for example a coworker pressures the interviewee to perform the interview instead of doing it himself. Another factor that Jacobsen (2002) brings up is to provide the interviewee with sufficient information about the research; its purpose, what pros and cons the study might bring for the interviewee and how the data is going to be used. While contacting the interviewees we have explained our subject and the purpose of our study. When they agree to perform an interview we have sent them some of the general questions and informed them of how the data will be used. This way the interviewee can make sure that he is the right person for the interview and that he has the right competence to answer our question based on practical experience.

The *demand on confidentiality and privacy* is another requirement that is just as important as the informed consent. It is important for us as interviewers to not encroach on the interviewees' personal space or that we disclose any sensitive information about the company. The interviews that we have conducted have not been focused on the specific person or the specific company that we are interviewing. Therefore we have not been interested in specific sensitive data. We have still been meticulous to not expose any data that the interviewee does not want exposed and to anonymize any data that could be sensitive. Since the transcripts have been sent to the interviewees we could make sure that we do not expose anything that they do not want exposed.

The last requirement is the *demand of an accurate review*. It is important that the collected data is correctly represented in transcriptions, that it is not taken out of context or misrepresented (Jacobsen 2002). To make sure that we have not represented the data in an inaccurate manner we have given the interviewees an opportunity to read through our transcriptions. The transcriptions are also put as appendix at the end of the thesis to ensure that the data can be put in its correct context.

4. Results

This part of the research paper will present the results. The interviews we performed covered two main areas; IT capabilities and Value creation of big data. Furthermore IT capabilities have been divided into two subheads shown in the theoretical model used; Competencies and Practices. The results will be presented within each associated area.

4.1 IT capabilities

4.1.1 Competencies

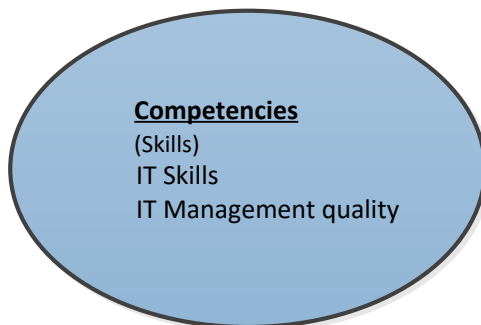


Figure 5, The competencies derived from the theoretical model.

From the data analyst's own experiences, the customers have described their data as black holes before implementing data visualization solutions. It was impossible for them to get a good overview of their company. The customers of the data analyst do not always trust the data either, they tell themselves that there must have been an error in the collection process. This shows how important IT competencies actually are. One of the examples we found most interesting from the data analyst was regarding a production company where the data analyst had created a data visualization application. In the application it was revealed that there were 100 000 performed tests on the products instead of the supposed 10 000. While analyzing the data they found out that the same tests were run over and over again in their China factory. By further analyzing the case it turned out that the executives in China got bigger bonuses in relation to how many tests they had performed so they had just copied the same test cases every week in order to receive a bigger bonus. Obviously this led to big process changes for that specific factory and helped the company save both time and money.

The ability to understand the data and knowing what to do with it requires strong teams with different types of expertise. The data analyst expresses that there are companies not being able to utilize the data they gather because of a lack in digital competence. The big data process looks different for different companies. The data analyst's regular customers describes their core business, processes and their problem and it is then up to the data analyst to understand what is important, what can be improved and what kind of data should be gathered. Afterwards it is up to the data analyst to create a relevant data model and the scripts needed to visualize the data. The data analyst from Softhouse expresses how the big data process would be optimized if the customer would have a better digital understanding and be able to ask the right questions from the start. He also expresses how every part of the process is equally important.

In regards to digital competence, the data analyst further explains that the GUI must be user friendly and intuitive. If a manual is needed for a business user to understand, then you have failed. Most often it comes down to if the users are experienced with the data visualization tool. If they are experienced there are usually no problems and they are able to look about on their own. Education is key he means, it builds up both the interest and the necessary competence to utilize the big data results. On the other hand, the tools that the business people use cannot do any real harm since they only display results, no manipulation of the collected data is involved. *"The dream user is well versed with his data and knows from the start the questions he want to have answered"* (Line 176, Appendix 3).

To interpret and draw conclusions from big data is a complex activity. As the person making the demands, it is important to have a clarity of what questions are going to be asked and what will be done with the answers, an interviewee claims. In order to know what data should be collected, strong knowledge of your company is needed. However one of the CEOs states, once you start collecting, you will learn continuously which data is interesting and which is not. From that point on, you start collecting and mining in a smart way. Avensia has learnt from years of experience with big data how to address different types of companies since their behavior is similar. Their digital competence has grown over time and is now at a point where they feel comfortable. However when they get a new customer with a rare product or business model, they feel that they have to learn together with the customer in order to build their website. Everything comes with experience he claims. *“Big things, common behaviors and big trends we know of (...) but with a changing environment and constantly new technologies we have to adapt and learn”* (Line 116, Appendix 5).

King do not have the manpower (IT capabilities) nor the tools (IT assets) to internally gather big data. Instead they take part of their customers' collected data. Many times they have to ask their customers about specific data. Other times the customers themselves send over the data they think is relevant. King will thereafter work together with the customer to find the important parameters to filter the data in order to gain insights. The interviewee states that the problem is that the data they get only gives them a look in real time and no real insights. King is still in a phase where they feel that they are too small to hire a data analyst. Instead for now, the planners are taking care of all the data processing. It takes strong cooperation between planners and the customers in order to make sense of the data. However the interviewee also says that even though it requires more digital competence to gain value from big data, he also believes that the tools for analyzing big data will get better and easier to use. *“Regarding production and analysis of digital data we have a lot more to learn and if there is anybody selling that kind of education it would be of great value to King.”* (Line 178, Appendix 6).

The interviewee does not think that King is in a position right now where they are able to hire a data analyst but they understand that it can be valuable in the near future. The interviewee believes that the digital competency that King has right now is enough since the amount of data they are working with at the moment is relatively small. But with time, the amount of data will increase and they will need an improved data competency.

The collection of these kinds of data is not free. That is one of the biggest reasons why King is not collecting on their own, the interviewee explains. King has a consumer panel consisting of 400 people which they retrieve data from. However this is nothing in comparison to the databases of for example ICA, containing data about every consumer who ever bought groceries.

A factor regarding IT competency that also needs to be accounted for is the structure of data. The data needs to be structured enough to be informative and organized enough to make the data searchable. The collected data may differ from each other and be structured in different ways. For example data collected from businesses in Norway and Sweden may both contain the same records but the data from Norway is structured into intervals while the data from Sweden is in absolute numbers. The interviewee explains the process of ensuring the quality of the data where they first conduct a couple of different machine tests on the data to make sure that it is imported correctly. They also have spot-checks and manual checks on the data to see if it looks reasonable. Having to manually check or structure the data puts demands on the competency of the employees. One little error might lead to dire consequences. Our interviewee brings up an example of a car insurance company that made a mass send out to all of their customers. They somehow managed to move one cell one line down. The mass send out resulted in people getting mail where none of the addresses corresponded with the identification numbers, all of that just because the cells had moved one line. Even though this example was based on a mistake it still shows the importance of having the right competency to handle the data.

Outsourcing

With the vast amount of data out there it might be hard for small businesses to make sense of the data on their own. It might instead be better to invest resources in external actors who can assist with picking out the right information and structure it. Let the external actor take notes of what questions that need answers and package it in a way that allows your company to make sense out of it. In the collection process, not much digital competency is required since most of it is automatic. However without the ability to analyze the collection, problems would obviously arise without the right digital competence.

4.1.2 Practices

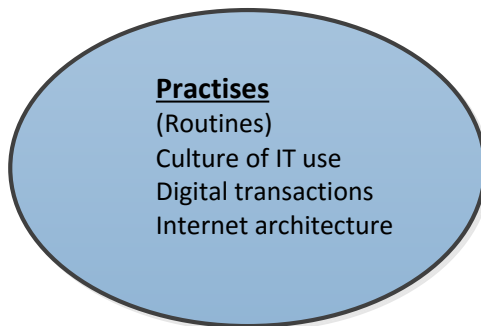


Figure 6, The practices derived from the theoretical model

An interviewee states that if the management team has to keep an open mind towards new trends then you got to have a culture that is assenting. A company cannot have a culture which treats every new trend as something interim that will be gone soon. Companies need a culture which is based on curiosity.

One of our interviewees further states that in order to utilize big data, a company needs a culture which is ready for change. If people feel uncomfortable with changes and have a negative attitude towards sudden changes, they will have a hard time working with big data. Furthermore the interviewee highlights the importance of understanding that changes are happening and that the new generations who are comfortable in this constant changing and fast-paced environment will have an advantage. There may also be organizations out there understanding that something has to be done to adapt to this new environment. But they have difficulties assimilating what is required from them. The interviewee from King believes that big data will continue to grow and that King are adjusting more and more into a more data-driven decision making culture to better respond to the growth of big data.

Another factor that has become increasingly important in order to utilize big data is that the manager has to be more "metrics driven". Today it is harder to base a decision on pure intuition or gut feeling. The interviewee from Lundalogik states: "*My job has gone from being very intuitive to being completely data driven*". (Line 257, Appendix 4). Companies utilizing big data, bases their decisions on analysis of the data where they can predict the outcomes of their investments. This might be something many businesses will learn the hard way.

According to one of the CEOs it has become increasingly important for managers to keep an open mind. It is further important to understand that if you are lacking the right competency yourself, then you have to procure the right resources and put faith in the personnel who make those decisions. Lundalogik has for instance over the last few years changed their job specification radically for their employees in the marketing department. Nowadays they hire data analysts, engineers and system scientists where they previously hired people who were creative and good at writing. They still keep the creative people, but they are supplemented by data analysts, who work full time with analyzing patterns from the data. A big part of their marketing is based on creating relevant content, which puts higher demand on producing valid content in a website for instance. A lot of the data collected is used to discover patterns and if you are in a manager position then it becomes of high importance that you can motivate your choices. You have to be

able to motivate hiring an engineer to work at the marketing department as a data analyst. This puts demand on an open minded management group so they can see the relevancy and the potential of added value from the appointment.

One of the CEOs explains how they respect integrity: In regards to personal integrity, the customers are in a way anonymous. Everything is collected through cookies so most information is just connected to an anonymous cookie. *“If the cookie has acted this way, then we should treat it like this other cookie with the same previous behavior”*. (Line 35, Appendix 5) Further there is always the availability to visit the website in incognito-mode where the user can stay completely anonymous. The CEO also explains the importance of having a culture of IT use that shows respect for personal integrity, you do not want the customer to feel that his privacy is violated in any way.

4.2 Value creation

Big corporations have always been able to afford to work with data analysis in order to find out where their frequency is and what potential they have. Smaller businesses on the other hand have always believed that it is too expensive, that they do not have time or that they do not believe that they need data analysis one of the CEOs claim. With today's big data tools this is not necessarily true, but there are too few businesses who realize the potential of the data the CEO tells us.

4.2.1 How big data creates value for organizations

Avensia were quick to introduce the big data way of working. In 2010-2011 Avensia were one of the first to work with personalization and collection of customer behavior which in their case led to competitive advantage. In the beginning their customers were skeptical about collecting this kind of data from customers and could not see the point of doing it. They were not able to understand the possibilities of segmenting websites on a personal level or from customer trends found. A classic segmentation is the statement that women in their fifties shop different than men in their thirties, the introduction of big data however made it possible to go down on personal levels. The new statement is that customers with similar behaviors are interested in similar products. Without enough data collection it is not possible to work on such personal levels. The CEO tells us that it was a hard task in the early stages of the big data era to convince customers why they should collect as much data as they are doing today.

The CEO of Avensia feels that using big data and collecting information about customers is a win-win situation. *“As a customer you want relevant information and the ability to find what you are looking for fast, I would be irritated if the site would forget about me.”* (Line 45, Appendix 5). From the business perspective, it is a huge advantage to be able to collect all this data. If the website is not able to show the right products the first times the customers visit, the customer will go somewhere else.

Another case that was presented to us shows how big data in the form of GPS monitors has led to new business opportunities. We were shown how a relatively new company had emerged with the business idea of putting GPS monitors on trains in order to see how they are moving. Since there might be cargo which is sensitive to certain conditions like becoming heated or chilled, the company also provides temperature monitors. Furthermore the monitors gather information regarding rough braking, collisions and other occurrences that should not happen. The company then provides the customer with all this data in real time and makes it possible for them to optimize routes and strategically place cargo etc.

“Big data helps us with our creative deliveries” (Line 98, Appendix 6). The interviewee from King tells us how they formed a tourist campaign based on gathered GPS-data for one of their customers. The customer which is a service-based company has gathered enormous amounts of data over the years of their cars driving patterns. King used this collected data and formed a campaign targeting tourists visiting Stockholm. By looking at the data, King found out where most pick-ups and drop-offs are. Thereafter they used this data to find out where people actually were heading and created a tourist guide based on where Stockholmers had been going, connected to google maps.

4.2.2 Value opportunities

The CEO of Avensia states that the big data possibilities are endless, but there are some problems in the way. Firstly they would want to collect data from all of their customers and in turn all of their customers' customers. However there are legal issues with collecting from a third party. Secondly they are interested in collecting aggregated information across client boundaries. *"This would help us maximize sales for our customers"*. (Line 136, Appendix 5). The issue is that the end users have not expressed consent with sharing their information to a third party. So there are legal issues stopping Avensia, but the CEO claims that there is enormous potential from collecting customer data as well as behavior data and that it really would drive Avensia forward.

He also believes that there are endless opportunities for e-commerce with big data. *"Today e-commerce is a strong salesman, however in the future, e-commerce will be stronger than any other human salesman"*(Line 167, Appendix 5). This is because e-commerce knows everything about the customer and can therefore present the most relevant products and treat the customer the way the customer wants to be treated. E-commerce will be taken further than any human salesman ever will be able to. The CEO of Avensia further claims that only 1% at most of the collected data is used to gain insights. *"It is extremely underused when you think of all the possible conclusions you can get from big data."* (Line 99, Appendix 5). The big players have sold big data solutions for a long time, but the companies still have a long way to go to figure out what to do with big data.

Our interviewee from Styleport considers the analysis of customer data as one of the biggest opportunities with big data. By analyzing data, the company can find patterns and connect them to time periods and thus anticipate upcoming trends. However, companies also need to look at what the customers do not purchase, what they are curious of but choose not to proceed with. The company could then try to get an understanding of what factor made the customer change their mind and create a solution for it.

For Styleport, the first step is to increase their knowledge about their current customers. The ones that already bought products or visited the website. They feel that they are still too small to consider working with big data. However the interviewee expresses the possibilities in the future, especially to improve the customer experience with more personalized help.

For King, it is common that there are complicated and complex databases which are not open for them. Instead the interviewee tells us that they have to ask their customers for certain parts of the data. He feels that King would greatly benefit from getting access to more data although more digital competency is probably needed.

4.2.3 Big data and marketing

The Deputy CEO tells us how Lundalogik works with analysis of all their in-flows of data. In order to find new prospects they try to analyze the user-behavior on their website by using cookies. They collect data on what pages the user has visited, when they did it and how much time was spent on each page. By analyzing these patterns they can cross reference the data with the collected company data to see if the user matches their customer profile. Another technology used to find new customers is machine learning. With machine learning they try to make predictions based on the patterns how probable it is that the user will become a customer or not. In the end all the parameters are put together into a value which states how hot the lead is.

Another way big data is used is in Programmatic marketing. The ability for companies to automatically market themselves online where their consumers are moving. The interviewee claims that it surely is highly effective if you want to spread information to a targeted market. The amount of available data is growing continuously and there are more insights to be found. At the same time there are more behaviors and hypotheses to confirm. The interviewee claims that big data has helped King to reach their goals in an earlier stage of the process and made it easier to address the correct target group.

5. Discussion

In this section we discuss and reflect upon the results we have acquired from our empirical research and link them to the theoretical framework presented in the literature review.

5.1 IT Capabilities

5.1.1 Competencies

Among the companies that participated in our study, there is a mutual conviction that there is a lack in digital competence in several companies and fields. They all agree that the lack reduce the ability to utilize big data. Regarding IT competency our results have shown us the importance of management being able to ask the right questions. Gao et al. (2015) calls this the involvement of “the right people”. As we got to interview different key players in the big data process we got to see the importance from different perspectives. It is important to have a clarity of what questions are going to be asked and what will be done with the answers. In order to know what data should be collected, strong knowledge of the company is indeed needed. Our interviewees believe in the creation of what Gao et al. (2015) calls “the right people” through education and experience. Avensia have created and gained relevant IT competencies through working with several similar customers and discovered certain trends and best practices. The importance of management asking the right questions was made especially clear from the data analyst as the big data process then would be optimized from the start.

In big data projects, we understood from the data analyst’s perspective that a strong team is required. One single data analyst can provide insights but to achieve more value from big data, a stronger team is required. As we gained this insight, we found it to be in line with what Sicular (2012) previously had found in her research. The strongest results are gained while having a multidisciplinary analysis team with a wide range of different expertises. According to Chen et al. (2012) the skills required include management of unstructured data, text mining, natural language processing, machine learning and social network analysis. However, data analysts are expensive and therefore a Return On Investment analysis might be in place. More than one data analyst might be beneficial depending on the size of the project.

For Smaller companies, the first step might be to increase their knowledge about their current customers. The ones that already bought products or visited the website. Perhaps they feel too small to consider working with big data. However the interviewee from Styleport expresses the possibilities in the future, especially to improve the customer experience with more personalized help.

5.1.2 Practices

Having a corporate culture that facilitates changes and keeps an open mind toward new trends is something that several of our interviewees have brought up. This is also something that Gao et al. (2015) brings up when they speak about big data projects. The authors also highlight that in order to work with big data, the project group needs to be flexible and quickly able to adapt and react to sudden changes in the environment. Another interviewee brings up the importance of organizations understanding that changes are happening, which further reinforces the theory of having a culture that keeps the employees comfortable with constant changes. The importance of having an open culture could be derived from the rapid development of new technologies and practices in the business environment. Big data is just one of these technologies but implementing a big data solution provides some major changes for companies.

Another trend amongst our interviewees seem to be the transaction from intuition based decision making to a more data driven decision making. A clear example of this is the marketing department of Lundalogik where they have gone from hiring creative people to instead use data analysts to market themselves. It seems that it is no longer sufficient to base your decisions on intuition alone if you want to stay

competitive. This statement is in line with Watson (2014) where he talks about the positive relationship between data-driven decision making, organizational performance and competitive position. With such vast amounts of data available it is hard to compete with competitors that is able to analyze their customers' behavior and better predict upcoming trends. Watson (2014) further reinforces that it is necessary for organizations to base their decision on facts in order to gain value from big data. To better utilize the facts gained from the data analytics, companies would benefit from a data driven decision-making culture. The Deputy CEO of Lundalogik further emphasizes the importance of being able to motivate decisions to stakeholders. Using big data solutions provides facts that further strengthen propositions. Big data solutions with a data driven decision making culture could thus lead to increased adaptability and competitive advantage.

Besides an open minded and data driven culture it could also be beneficial for companies to have a culture that facilitates a respectful approach to integrity and ethics. Given all the data that is collected, there is always a chance for infringement on people's personal integrity. Therefore it is important for companies to work on their culture of IT use and how data is handled as to not get into legal trouble or accidentally expose sensitive information. Sicular (2012) argues about the importance of including legal experts in big data projects from the start. Having a clearly defined legal framework from the start may allow organizations to avoid hassle or lost profit through ethical, security or legal issues regarding big data.

In Aral and Weill's (2014) theoretical model of IT resources displayed in figure 4, there are three different types of practices presented; *Culture of IT use*, *Digital transactions* and *Internet architecture*. During our empirical research we have found that our interviewees almost solely talk about the culture of the company as a practice needed to create value from big data. This might be because most of our interviewees have been CEOs and as a CEO you tend to have a holistic approach to running the business. Seeing the organization as an entity where you have to get all the employees working in the same direction towards a common goal, the culture becomes a core practice. The systems for digital transactions and internet architecture are material resources which can be procured as long as there is financial support for it. The culture is a more intricate matter which cannot be procured through financial means which might be why the CEOs emphasize the importance of having the right culture in order to create value from big data.

Aligning the management and organizational culture in order to leverage from new insights gained is something that Gao et al. (2015) and Aral and Weill (2007) remarks on. According to them, this will help the organization to utilize the full potential of big data and we find this to be in line with our findings. A constantly changing environment further requires an alignment of several business areas in order to stay flexible and ready for changes based on insights gained.

5.3 Value creation

All interviewees agree that using big data solutions creates win-win situations, meaning that it is beneficial for the company as well as the customers. However one of the interviewees mentions that there are extreme cases where the customer might become offended. All companies interviewed are also expressing future opportunities with big data. Watson (2014) claims that companies need a proper alignment of technology, people and processes in order to succeed with big data. The interviewees seem to understand the importance of alignment, the best example being Lundalogik as they are now single handedly hiring data analysts to their marketing division. Another important factor to note in regards to alignment is business culture. A trend found from our studies is that all interviewees speak of a data driven decision making culture, even though they might not use the exact same term for it.

We also found out from our studies that companies have a hard time utilizing the full potential of big data. The CEO of Avensia believes that only 1% of all data that is collected is used to gain insights, meaning

that it is extremely underused. Smaller businesses find big data projects expensive but one of our interviewees believes that the smaller businesses find it expensive because they are not able to see the return on investment. This is in line with Watson's (2014) opinion that the business users don't need full understanding of all the technology but they do need to understand the possibilities of big data. We believe that this is connected to a lack of digital competence and being able to understand what big data actually enables for companies.

From our studies we tend to agree with the claim that *"big data is the most transformative paradigm to hit sales and marketing teams since the advent of the telephone"* (Hollison, 2015, p. 1). We can see that companies are moving into a new decision making culture which involves a trust of collected data. By looking at current customers, companies are able to identify value creating information. With the use of big data, managers are now able to understand who their customers really are. Everything from what they usually eat, what clothes they like to wear and where they like to go on vacation is now collectable. By understanding this, companies are better able to customize their products to fulfill the needs of their paying customers. Furthermore new market segments become targetable and therefore enabling a larger customer base through smart marketing. Online marketing simplifies targeting the right demography. Age, area and similar interests are key values that now are selectable for the marketer.

To really understand the customer's needs and behavior is a necessity in order to create a good product. A solid target group analysis is a good basis for creating a good product, big data can be used and analyzed to create the target group analysis. Even if there are clear goals defined by the business from perhaps effect mapping, the product will not be effective if it does not meet the target group's needs.

6. Conclusion

Having the Resource Based view applied as proposed, IT capabilities are defined as an important correlation between competencies and practices. IT capabilities do not solely create sustained competitive advantage. Instead they work as a complementary and can be seen as a piece of the puzzle showing a company's resources.

The research question that was initially presented in this thesis was;

- *What IT capabilities does a company need in order to create value from big data?*

Our empirical research shows that there is a new culture arising with the big data era. Decision makers are moving into a more data-driven decision making culture. Big data is complex especially in accordance with gaining insights and value. Strong teams with different sets of specialties are required in order to fully utilize and gain value from big data. Furthermore it is important that the people involved are flexible and able to trust the data.

Buying a product today involves sharing intimate details regarding consumption patterns with retailers. A large part of the retailers will use this information in regards to optimizing their products, stores, customized offers or target segment. For businesses to thrive in the future, we recommend managers to learn how big data can create value. A strong technology competency is not necessarily required, however an understanding of the value creation is important to stay competitive. Data scientists are employed for this exact reason, finding valuable strategic insights having the required and relevant IT competency. Furthermore it is important to understand that just collecting data does not create any value in itself, the value can be created through proper analytics and data-driven decision making.

Big data has enabled personalization with the broad collection of customer behavior. Classic customer segmentation, meaning the ability to state that one specific group of people shop differently than another group is now being replaced by personalization. Big data is making personalization possible, meaning that the customer will always get a unique experience based on his or her preferences. Furthermore, big data is now enabling the creation of the strongest salesman of the future, e-commerce. The statement is quite simple really; customers with similar behaviors are interested in similar products.

Another important factor is the data driven decision making culture. However without the right competencies from decision makers, the data could easily be misinterpreted and conclusions could be wrongfully derived. In order to gain relevant competencies and being able to achieve a successful culture, experience and education is key.

Ways of collecting data are expanding continuously. From our research we believe that some companies, perhaps with more traditional forms of management, do not understand how insights gained from social media, internet habits and other collectable data can create value for a business. The ability to use this complex data in order to create value is a big potential for businesses worldwide.

7. Appendix

Appendix 1 - Interview Guide

Background

- Role and background of interviewee
- Current work with big data
- Reason for implementing a big data solution
- How their work has changed with the big data era

IT capabilities

- What IT competencies are required from a decision maker / analyst to utilize big data
- What the big data process looks like
- Requirements for answering relevant questions in relation to data driven decision making
- The required IT competencies for an organization to gain value from big data
- How the corporate culture affects the work with big data

Value creation

- How big data creates value for their organization
- If and how they use big data for marketing purposes
- If big data has led to increased customer satisfaction
- What opportunities has big data enabled for decision makers
- The correlation between IT competence and value creation of big data

IT assets

- The tools required to gain insights from big data
- Importance of data visualization

Appendix 2 – Styleport

Interview by: Phone

Position: CEO

Date: 2016-05-09

M = Mattias

V = Viktor

- **Start of interview**

1 M: Berätta om din arbetsroll

2 Min arbetsroll är att jag är både ägare och ansvarig för företaget. Vi är ett litet företag och jag gör det
3 mesta här. Vi är än så länge pyttesmå och håller på att starta igång.

4

5 M: Hur vill du se att ni ska jobba med big data i framtiden för Styleport?

6 Vi är fortfarande en... Jag hoppas att vi kommer växa och bli så stora att vi kan dra nytta av big data fullt
7 ut. Men jag tror att det första vi behöver göra är att lära oss mer om våra egna kunder, de som redan finns
8 eller de som är intresserade, de som kommer in på vår site. Utvidga den informationen på ett effektivt sätt.
9 Återigen, vi har inte så mycket volym än så länge så att vi kan dra stora växlar utan jag tror det handlar
10 mer om att hantera de kunderna vi har, de som är intresserade av det vi har att erbjuda.

11

12 V: Hur kollar ni på de kunderna ni har? Tar ni in någon form av data?

13 Det är lite som jag sa till dig i onsdags, att det handlar om att vi har köpt det här företaget för ett halvår
14 sedan eller sådär. Det har varit så att det företaget handlat på ett visst sätt och jag säger att vi behöver göra
15 om... Och vi håller på att ställa om nu. Bland annat så är en del i det att bli bättre på att utnyttja
16 informationen och kunskapen om våra kunder så att kunderna får mer personlig assistans än bara generell
17 hjälp.

18

**19 M: Så just nu får ni egentligen inte in någon som helst information om vilka som handlar på er site
20 eller?**

21 Nej. Alltså vi kan ju se vem som... Telefonnummer och mailadress och så vidare och vi skulle kunna
22 utnyttja till exempel mailadress för att skicka nyhetsbrev men vi har gjort det sporadiskt men vi har inte
23 någon metod idag som känns tillfredsställande.

24

25 M: Vad känner du just nu då? Vad är viktigast att samla in om era kunder?

26 Det viktigaste att samla in, det är dels vad de köper naturligtvis och försöka se ett mönster ur det och
27 koppla det till tidsperioder så att man över tid ser ett mönster, av vad de är intresserade av. Men sen också
28 vad de inte köper, vad de är nyfikna att titta på men sen väljer att inte... Varken lägga i kassan eller avsluta
29 köpet.

30

31 M: Lite angående marknadsföring, hur marknadsför sig Styleport?

32 Återigen jag kopplar tillbaka till var vi är någonstans. Det finns ett nuläges och ett slags framtida läge. Så
33 om vi säger nu då så utnyttjar vi oss av Prisjakt och Price runner. Två kanaler för att synas och sen tror vi
34 att våra kunder är mycket inne på Facebook. Vi har inte kommit igång, vi har gjort lite på Instagram, men
35 inte speciellt mycket. Vi har inte varit aktiva på Google heller, det är ren 5:02???) uppsikt vi har liksom.
36 Lyckas fånga upp saker på vissa av våra produkter. Det vi gör, för att sammanfatta det är:
37 Facebook,Prisjakt och Pricerunner i dagsläget.

38

39 V: Vad är nästa steg för er?

40 Nästa steg är Google adworks tror jag och så kanske någon typ av marknadsföring eller organisk
41 marknadsföring. Jobba bättre med SEO och bli duktigare på det. Och sen också utnyttja kanalerna,
42 Facebook och Instagram och skapa en bättre känsla på de kanalerna. Inte bara genom att göra rena avslut,
43 utveckla våra egna konton på Facebook och Instagram så att det är relevant för våra kunder att följa oss.

44

45 M: Kör ni några ads på Facebook också eller är det bara att ni har era kanaler?

46 Litegrann har vi gjort det, det kommer vi fortsätta med. I relation till pengarna så får vi bra trafik, vi pratar
47 jättemycket avslut när vi får bra trafik. Ibland så har man ingen trafik och då får man inga avslut heller så
48 att det gäller att skaffa sig trafiken. Se till så att man har så bra grejer så att folk köper.

49

**50 M: Om jag förstått det rätt, när man väl lägger upp en Facebook-ad så bestämmer man väl exakt
51 vilken marknad man vill rikta sig mot?**

52 Om jag säger såhär. Man tror att man styr ganska väl men sen har man ingen aning om vad som händer
53 bakom kulisserna. Man kan välja ut kundkategorier utifrån intresse och ålder/kön,
54

55 **M: Hur gjordes den analysen för att bestämma det? Vilken grupp man vill rikta sig mot?**

56 Just det, precis. Man kan segmentera på ett ganska bra sätt och det är en jättestyrka från Facebook att man
57 kan det. Om man jämför med annan digital media som kommer från alla digitala tidningar och... Där
58 skjuter man rätt brett och behandlar annonsering på tidningar, vilket jag är mer tveksam till. Den storleken
59 av företag som vi är, är inte tillräckligt stora för att göra generell marknadsföring av... såhär
60 varumärkesvidgande. Det är ett ganska dyrt sätt att annonsera.

61

62 **M: Lite det jag vill komma åt är väl hur ni gjorde den här analysen för att förstå vilka kunder ni
63 ville rikta er mot?**

64 Bra fråga. Om vi tar koll på vilka våra kunder är... Det är liksom ometodisk analys eller någon slags
65 känsla av vad som mest sannolikt borde funka, skulle jag vilja säga utifrån erfarenhet.

66

67 **V: Vad krävs det för kompetenser tror du? För att kunna utnyttja de här big data- lösningarna
68 inom företaget?**

69 Det är en otroligt bra fråga. Jag tror att... Nu gissar jag lite grovt här. Jag är inte jättehaj på detta men jag
70 tror att det är relativt komplext för att kunna förstå och kunna lära sig att ta ut någonting av den
71 informationen. Dels måste man som kravställare ha det klart för sig vad man egentligen vill komma åt och
72 vad den informationen man får ut, vad man ska göra med den. Det finns ju sjukt mycket information att
73 plocka ut, men sen måste man också kunna tolka den och då tror jag att man behöver hjälp av... Som liten
74 till medelstor aktör så kan det vara svårt att lösa det själv utan man måste ha hjälp utav någon extern part
75 som är duktig på att plocka ut information, strukturera upp den. Anteckna de frågor som man behöver ha
76 svar på och sen paketera den på ett sätt så att man kan få ut något vettigt av den. Jag vet inte om det var ett
77 svar på din fråga men jag kan inte den mer i detalj vad som krävs för kompetenser.

78

79 **V: Kommer ni då ha några verktyg internt eller tänker ni att det får vara en extern part som sköter
80 allt?**

81 Initialt i alla fall, att hantera vår egen data och information som vi har. Och sen också om man vill gå
82 externt och försöka förstå vad det finns för trender och kundbeteenden och så.

83

84 **P: Alltså får jag ställa er en fråga, vad är er bild av... Vad händer med big data, har ni någon tes kring det
85 här?**

86

87 **M: Det vi ser och det som de flesta akademiker säger är att det här något som växer och växer, det
88 kommer bara bli större. Men det vi kollar på egentligen är väl lite från ur ett företagsperspektiv
89 vad som krävs för kompetens, vilken kultur och lite sådana faktorer vad det är som gör att man
90 faktiskt kan lyckas med den här datainsamlingen och dra nytta av den. Det vi tycker är mest
91 intressant är ju till exempel på marknadsavdelningen, att man nu längre inte anställer till exempel
92 kreativa människor som ska skriva reklamjinglar och så vidare. Nu anställer man faktiskt IT
93 människor eftersom det är de som ska göra den här analysen och förstå vem kunden är och utifrån
94 det göra riktad marknadsföring och så vidare. Det är väl lite det vi har tittat på.**

95 Jag förstår. Och ni vill förstå litegrann hur företag idag, var de står någonstans. Olika storlekar på företag
96 då också antar jag.

97

98 **M: Ja och det är väldigt mycket prat om, framförallt i Dagens Industri att det är brist på digital
99 kompetens bland de här svenska företagen. Att vi ligger väldigt mycket efter på den digitala sidan
100 jämfört med många andra länder. Då kollar vi på hur pass viktigt det faktiskt är med IT kompetens
101 i företag.**

102 Samtidigt så är det väl lite motsägelsefullt att Sverige har en väldigt stor andel e-handelsföretag jämfört
103 med många andra länder. Någoting finns det väl som kunderna... Som företagen gör, att kunderna på
104 något sätt uppskattar när det finns otroligt mycket mer att förstå kring kundernas beteende.

105

106 **V: Ja precis. Men tror du att företagskulturen har någon påverkan på dataarbetet. Måste man ha**
107 **någon speciell företagskultur eller vad tror du?**

108 Absolut. Relevant fråga. Man måste nog ha en förståelse där man... Jag har ju varit med om båda sidorna,
109 både sälj och det digitala. I alla fall ur ett handelsperspektiv så har jag jobbat både innan den digitala
110 revolutionen och efter. Och jag tror att de som... Finns de dem som känner sig obekväma i det som händer
111 och det påverkar deras beteende. Då tror jag att om man inte är med där och förstår att man måste vara
112 med och skifta och ta in en yngre kompetens som mer är vana vid att leva i den här miljön. Då tror jag att,
113 det blir knepigare då. Det finns nog exempel på det att man förstår att man behöver göra någonting men
114 det är svårt liksom att ta in och ta till sig vad som krävs.

115

116 **M: Lite det vi har fått höra också är väl att vissa företag har lite mer negativ syn på det hela och ser**
117 **det mer som en fluga ungefär.**

118 Ja då är man nog halt ute om man ser det som en fluga. Ja, men lite så. Men det kan nog vara så att man
119 tycker att... Sen ska man ju ha någon slags sund skepsis även där. Man behöver inte vara först och springa
120 på alla nya bollar som snurras. Att man kan läsa ur ett mikro... Ha mycket nytta av digital information
121 eftersom du har all information om hur kunderna beter sig och också kan trendanalysera och se hur
122 kunderna också kommer bete sig. Att man ser vad som skrivs i det digitala och sådana grejer.

Appendix 3 – Softhouse

Interview by: Phone and shared screen

Position: Data analyst

Date: 2016-04-14

M = Mattias

V = Viktor

-
- **Start of interview**

1 1:28 **V: Berätta lite om din arbetsroll.**

2 Jag jobbar med QlikView åt ett gäng olika kunder, stora och små och visualiserar data ur deras system. Styrkan där
3 är att med QlikView kan man koppla ihop datan från olika system så du kan koppla ihop från arbetsrådssystem
4 tidsskrivningssystem, ekonomisystem, kundrelationssystem och få en helhet, en överblick över alltihopa.

5

6 **M: Just det. Intressant, det är ju perfekt.**

7

8 **V: Hur gör ni då, hur arbetar ni aktivt med den datan ni får in?**

9 Rent praktiskt hur det går till?

10

11 **V: Ja.**

12 Då kanske det är bra att jag har min skärm här i alla fall då. Det går till så att om man... Vi kopplar oss åt databaser
13 och använder Qlikview för att koppla oss mot databaserna. Så läser vi ut data ur tabellerna som finns i de här
14 databaserna då. Det kan vara ett gäng olika tabeller. Om vi öppnar en så sparar jag det i qwd-filer som är ett format
15 som Qlikview läser in väldigt snabbt och det är komprimerat så det tar väldigt lite plats och är extremt snabbt att läsa
16 in. För det är så om du ska använda samma data på fem olika ställen, du vill ha någon ekonomisk överblick där du
17 vill ha datan och göra någonting med den och så vill du ha en operatör eller någonting där han vill kunna se sin data
18 och kanske se hur många case det har varit. Då istället för att gå och hämta datan två gånger i systemet så läser jag ut
19 den en gång och så används den filen två gånger och då med slutapplikationen. Med Qlikview läser vi inte direkt från
20 databaserna, eller man kan det om man då ska köra riktig big data, då kan man inte läsa ut och spara det i filer. Jag
21 vet inte riktigt vad erat exjobb går ut på och vad det är ni håller på med?

22

23 4:24 **M: Det vi vill komma åt är väl egentligen: hur beslutsfattande förändras genom hela big data eran eller
24 vad man ska säga. Hur man enklare kan ta beslut utifrån all den datan man samlar in.**

25 Ja just det. Big data för mig, det är för mig... Alltså jag håller på med data, väldigt mycket data och det är väldigt
26 svårt... det går inte för en enda människa att titta i alla de här tabellerna och dra en slutsats. Men jag tycker inte att
27 det är big data för big data för mig det är att datan inte går att hantera på en dator om vi säger så. Du måste klusta, du
28 måste ha ett helt gäng datorer. Du måste ha det i molnet och ha tillgång till tusentals datorer eller ett tiotal
29 åtminstone. Då tycker jag man börjar prata big data. Men det här när jag kan läsa ut tabeller och spara dem i filer.
30 Visserligen har jag tabeller här som är sparade även i Qlikviews komprimerade format så kan de bli 3 gigg stora.
31 Men jag tycker att det går fortfarande att hantera på en vanlig pc. Men just beslutsfattande och så där absolut så var
32 vi på rätt område, det tycker jag att jag håller på med.

33

34 [Delad skärm] Men jag öppnar min qwd-fil nu då och det ser ut som en databastabell, det är bara rakt upp och ner,
35 ser prick likadant ut som det gjorde i en databas som om man hade öppnat i t.ex. Oracle. Utifrån de här qwd-filerna
36 som ligger sparade, i det här fallet som vi tittat på så motsvarar varje folder en databas. Här har vi deras databas för
37 budget och prognoser, ett speciellt system de har montörer, detta är för [Företag]. Ni har sett deras produkter, de
38 jobbar som underleverantörer till [Företag]. Om det är en produkt som inte funkar eller om det är någon som ska
39 flytta så får de ett meddelande på sin bärbara enhet och så kör den här service-killen på det ärendet då och så
40 registrerar han "det tog så här lång tid och jag fick göra detta och detta". Sen flyger det in i... Ifrån
41 ärendehanteringssystemet så kommer det bli en post i ekonomisystem som är det som är det som är här då [visar på
42 skärm]. Alla deras projekt och deras ID och vem som har ansvar och när de startade och när de slutade och olika
43 indelningar av projektet och så vidare. Här har vi tidsrapportering där de har skrivit tid. Lime för kundrelationer och
44 CRM-system. Här samlar de in primärdata från sina produkter.

45

46 7:41 **M: Får jag fråga hur det ser ut med kundrelationer, vad är det man samlar in då?**

47 Lime är en programvara som tillverkas i Lund.

48

49 **M: Ja från Lundalogik.**

50 När en kund har ringt så skriver du in här: Jag pratade med kunden idag, han var glad och allt var frid och fröjd och
51 sen har du en säljare knuten till varje kund och det är dem som sköter om datan och skriver in kundbesök eller alla
52 relationer du har haft med kunden. Behöver du bokföra något så var det ju... överallt du är och delar med dig av dina
53 uppgifter så kan du ge dig tusen på att du ligger i de här systemen.

54

55 8:32 **M: Det är lite det vi vill komma åt också, hur man kan hitta mönster och trender genom att få någon
56 slags överblick över den här datan. Och just det här med kundrelationer tycker vi är väldigt intressant i och**

57 **med att man då får en bättre bild över hur ens kunder fungerar, hur de tänker och vad de kanske har för**
58 **andra intressen och så vidare.**

59 Om jag bara går vidare här. Så här har vi tre databaser med mötdata från ungefär en halvmiljon elmätare då. (Namn
60 på systemet)9:11 är deras inköpssystem där de köper in prylar och varor. Utifrån detta då så gör vi olika typer av
61 Qlikapplikationer, det kan vara för deras ärendehanteringssystemet. Mätvärdena, HR- och inköpsavdelning, kvalitet
62 och processer, sälj, lite olika då.

63
64 Utifrån det här ekonomisystemet... Här kan man gå in och titta om man vill se affärsområde eller per bolag, de har ett
65 helt gäng olika bolag eller så kan man gå in på en enskild avdelning då. Kan t.ex. ta Skåne så får vi reda på direkt då
66 alla affärer som de har haft där då och så får jag se här uppe att de finns inte längre, de har antagligen blivit omdöpta,
67 eftersom det bara är 2014/2015 som är vita här. Vitt är associerat med det valet du gjort och grönt står för det valet
68 som jag aktivt har gjort. Allt som är vitt är möjligt att välja nu, det är okej att bryta ner det ytterligare. De här
69 leveransansvariga de jobbar på den här resultatenheten eller den avdelningen i Skåne, men så ser jag direkt här att
70 2016 har jag ingen data för. Vilket kan bero på att jag sitter med månad 3 här och att jag sitter med gammal data.
71 Men om vi tar t.ex. nordvästra Skåne så får vi med 2016 här uppe. Vi tar februaridatan och där kan vi se alla affärer
72 eller projekt som de hade igång i februari. Det finns olika då, kravning, fack, underhållsavtal och bla bla... hur
73 mycket som helst. Vilka kunder de haft då som har beställt jobbet. Nu tittar jag på en avdelning en månad. Här kan vi
74 se hur mycket man har sålt för på den avdelningen, hur mycket den direkta kostnaden är, lönekostnad. Indirekta
75 kostnader, skriva mobiltelefonabonnemang och så vidare och hur mycket som blir kvar i vinst. Här ser man då
76 februari 2014, och sen minskar det, här har vi marginalen och hur stor andel som är vinst.

77

78 **13:41 V: Hur ser då beslutsprocessen ut utifrån datan. De går in och kollar och ser det här, vad blir liksom**
79 **nästa steg?**

80 Det här är mest en uppföljning. Det här är för economiccontrollers då för att följa upp. Om vi går in i en månad här
81 och... Eller nej vi ska nog byta. Vi tar ett annat exempel. Den här var mer för att se "hur gick det", vem ligger på sin
82 budget och vem ligger under. Men vi letar upp ett annat fall där man kan dra lite slutsatser av datan.

83

84 **14:44 M: Funkar det då så att VD:n då ställer frågan: okej, hur gick det för den här koncernen. Och då är en**
85 **det en analytiker då som i sin tur tar fram de här resultaten eller hur fungerar det?**

86 Ja precis där har man en... Varje (karta???)15:05 görs det en ekonomisk rapport, hur går vi? Den sätts ihop varje
87 månad och då får man den här som kvitto, de här skrifterna och då sitter ekonomiska controllers eller
88 redovisningsbranscher då som går igenom och ser att ja men här har de ju inga intäkter alls, vad är det som är tokigt
89 här, är det fel inskrivet eller något annat som är fel. Och så följer de upp, vi har inte vart inne och ändrat eller
90 uppdaterat prognoserna efter senaste utfallet, så som det har blivit de har kontroll på verksamheten och använder
91 Qlikview då och all den här datan då för att bilda sig en verklighetsbaserad bild på hur det ligger till.

92

93 Men om vi kan ta ett annat exempel på... Kanske är lite lättare och roligare att förstå. Här har vi ett helt annat bolag.
94 Det här är ett bolag som är sprunget ur [Företag]. De var några grabbar där som fick en affärsidé: tågagnar sätter de
95 GPS-enheter på och så åker de här tågagnarna runt i Sverige och Europa och världen. Affärsidén då är att, bygger
96 på siffror på att 75-80% av alla tågagnar går tomma, de är på väg tillbaka efter de har lossat.

97

98 **V: Okej.**

99 Järnvägarna är översållade med tåg och det är därför persontrafiken inte kommer fram i tid när frakttågen är ute och
100 rullar och då är 80% av vagnarna tomma. Sen är det så att [Företag], de får en körning till Milano och då säger de att
101 "vi kopplar på två tomma vagnar och så får de följa med ner till Milano, det är bra att ha där eftersom det är en stor
102 järnvägsknutpunkt och rätt vad det är så är det någon som ska ha hem något från Italien så då har vi två vagnar,
103 skitbra." Sen är det ingen som ska ha de här vagnarna på ett bra tag och då glöms det bort att man har skickat dit
104 dem, så står det två vagnar i Milano. En gång varje år har de inventering och då ringer de runt till alla stationer i hela
105 Europa och hör om de har några vagnar hos dem. Det låter helt absurt men tydligen är det så det går till.

106

107 Så det de här gör är att de sätter GPS-doser på tågagnarna och de samlar även då in uppgifter om vilken temperatur
108 det är i vagnen eftersom det kan finnas gods i den som inte får frysa. De samlar in information om det har varit
109 någon hård inbromsning eller sammanstötning, om det har hänt någonting under resan och framförallt om vi är i tid
110 eller inte då. Så utifrån detta så finns det, sitter deras kunder och får tillgång till den Qlikview analysen och så sitter
111 de och analyserar var är våra vagnar mest och hur kan vi optimera flödet mer då. De berättade till exempel, ni ser en

112 kund som de har i Holland tror jag det var eller om det var i Belgien. De köpte in 140 stycken doser eftersom de hade
113 140 vagnar. Men de ska gå ner till 80 för de har kunnat rationalisera bort så många så att de behövdes inte längre.
114

115 Här kan man då vända och vrida informationen [delad skärm]. Här kan vi se att under januari 2016 så har vi fått in
116 438 positioner i Nederländerna och i Tyskland och så vidare då och därifrån kan man då välja de här två och bryta
117 ner det, i vilka regioner har vi varit och vilka stationer. Då ser man då att ja det är den där stationen som vi är i mest
118 och så kan man här direkt gå vidare där och se om man vill analysera hur mycket stilltid har vi haft då. Här en vagn
119 då och här är en vagn som har stått stilla då i fem dygn. Om vi då tar väck stationen så kan vi stilleståndstid. Här har
120 vi en vagn som stått still i 27 dagar på ett ställe och då är frågan varför och det beror lite på om vi då är en kund som
121 hyr ut vagnar till en annan då tycker man ju att det här skitbra, de hyr en vagn och så står den stilla. Men om jag
122 däremot är en kund som ska köra grejer då är det inte alls bra att min vagn inte utnyttjas, att den står och sover
123 någonstans. Så det beror helt på vilket fall man har. Sen är detta då sorterat på fallande ordning så här är några
124 vagnar som inte står stilla alls. Vi kan även göra en sammanslagen syn, hur mycket har de stått stilla totalt då i
125 januari 2016. Det här för varje vagn då egentligen och här har vi då 30 dagar på några vagnar, så de har stått stilla
126 nästan hela månaden. Här kan vi också sortera detta på något annat. Jag har inget case liksom förberett.
127

128 **22:35 M: Nej men det här är jättebra.**

129 Något de beställde här på sistone var att de ville ha (ships???) 22:44 och det är alltså... man vill veta hur lång tid det
130 tar om man har en kund som man vill köra från a, en hamn där det kommer in gods då och så ska du köra din
131 järnvägsvagn till kundens destination och så ska du lasta av där och sen vänder kanske vagnen och så ska du köra
132 tillbaka igen. Då vill du veta hur lång tid tar den här resan i snitt, hur lång tid tar resan, jo den tar alltid två dygn.
133 Men här är några fall där den tagit fyra dygn och varför gör den det, då blir det försenat. De har en annan kund
134 förresten som är [Företag] och de är lite roliga för de har fabriker i fyra olika städer i Europa. De har inga
135 lagerbyggnader på företaget, deras lager det är på tågvagarna, det är ute och rullar. Och i en stad då, så görs några
136 av produktdelarna, de görs i en stad och så lastas de på tågagnar och så görs vad det nu är i [stad] och så görs en del
137 i [stad] och så rullar de på natten. När morgonskiftet drar igång på i [stad] då ska grejerna vara där, det är just-in-
138 time. Är de inte där så är det ett helt skift som inte kan gå igång och börja tillverka grejer. Det måste vara på plats,
139 det måste vara i rätt tid.
140

141 **24:30 V: Vad gör man då om det är snöstorm eller så då?**

142 Då har man väl något kontrakt skrivet misstänker jag och så finns det avtalat hur mycket böter det blir om det är
143 försenat.
144

145 **24:54 V: Finns det någon felmarginal på datan, hur är datakvalitén, är den alltid 100 procentig eller är det så
146 att finns en risk att datan misstolkas.**

147 Ja, risken finns alltid att det är någon enhet som flippar ur och väljer att inte skicka meddelande eller bombarderar
148 med meddelanden. Det behöver alltid sitta i organisationen, folk som känner sin data, det är extremt viktigt att känna
149 sin data och att förstå helheten och veta liksom att här är en som har skickat en miljon meddelanden på två timmar så
150 då måste det vara enheten som flippat ur. Man måste känna sin data, det genomsyrar allt.
151

152 Om du väljer någon här, när (???)25:54 ska göras, det är något service ställe då. Där har de då avtalat att den får bara
153 stå stilla där i ett visst antal timmar. Längre än så får de inte ha den inne för månadservice, eller halvårsservice eller
154 vad det nu är utan då får den här serviceverkstan börja betala för att de inte har gjort sitt jobb i tid.
155

156 Här ser vi då det ljusblå, det är vägen dit och det orangea är vägen där det varit stilla. Det vi ser här är en enda resa.
157

158 **27:04 V: Det känns som att digital kompetens blir väldigt viktigt när man sitter och sysslar med det här. Finns
159 det företag då som sitter och inte riktigt vet hur de ska hantera datan?**

160 Ja, så kan det ju bli absolut.
161

162 **27:34 M: Har du haft någon direkt kontakt med någon slags beslutstagare och hört hur deras arbete har
163 förändrats sen de införde någon sådan här big data lösning?**

164 Mitt jobb består av tre delar, det ena är att sitta ner med kunden, ofta någon typ av chef eller ledning. De beskriver
165 liksom deras process eller affär eller problem och det gäller då att sätta sig in i den typen av problem och förstå vad
166 är det de håller på med och vad är det som är intressant här. Vad är det som driver kostnaderna eller vad är det som
167 de kan finlira. Sätta sig in kundens område och förstå det. Där också kommer då... Nästa tredjedel är då att kunna

168 sätta sig och göra det här skriptet som ligger bakom här. Bygga upp en datamodell som håller för att svara på de
169 frågor som man kommit fram till behövs frågas då. Så här ser datamodellen för denna applikationen som vi tittar på
170 här då. Här ligger observationerna i denna tabellen och så är det påkopplat en massa information. Den sista
171 tredjedelen är att kunna visualisera det för de som ska använda det, så att de kan få ut kraften ur datan och fatta rätt
172 beslut. De tre tredjedelarna är precis lika viktiga alla tre, ingen av dem har råd att gå sämre på än någon annan och
173 ingen man kan lägga mindre krut på en andra, alla är lika viktiga.

174

175 **29:40 M: Förstår de vilka frågor de vill ställa eller är det ofta ni som presenterar data för dem?**

176 Drömanvändaren är väl insatt i sin data och han vet vad det är han vill ha svar på och vad det är han är intresserad av
177 och varje gång han använder Qlikview eller han hör ofta hör av sig och säger "det här är ju skitbra, men vet du vad?
178 Tänk om vi kunde få svar på denna frågan också, kan vi lägga till det här, finns det möjlighet till att på något sätt
179 svara på de här frågorna?" Det är då det är riktigt kul och det är då det är ett momentum i det när det är användaren
180 som driver det framåt. Det blir inte lika bra att när det är jag som förklarar för dem, "Ja men här kan du se detta och
181 detta och det här är intressant." Det måste vara en väldigt intresserad användare eller (???) 30:42-användare. Där
182 måste man hamna att man gör han såpass intresserad av data eller att man gör det så intressant eller lättförståeligt att
183 det blir... att de nappar på det och bygger vidare på det. Det måste hela tiden förändras och förbättras, annars dör det
184 ut.

185

186 Nu gick jag in här på... drillade mig ner och tog den här stapeln mellan Europort, det är en jättestor, kanske Europas
187 största hamn till och med i Holland och så någon i (???)31:17. Jag väljer den resan och så ser jag alla vagnar som har
188 gått dit då. Så har jag valt en destination med många resor utförda. Då kan jag se här att jag har ett snitt här men det
189 är några som sticker ut och vad var det som hände just den dagen för denna vagnen. Nu har jag hittat en resa från en
190 destination som jag är intresserad av och jag kan analysera en enda resa, en enda vagn är jag på och så får jag ut all
191 information som tillhör den resan och så kan jag se här, vart var det han stod stilla. Han stod stilla här i ett dygn i
192 Rosehaven och ett dygn till, ja tre dygn till och med. Varför var det så och vad gjorde vi där. Så där kommer ju
193 domänkunskapen in, det måste man känna till som användare, vad är det som har hänt, är det något konstigt eller är
194 det normalt och hur kan vi göra för att bli av med det här glappet så att vi inte behöver stå stilla i tre dagar. Jag kan
195 visa andra exempel också.

196

197 **32:45 M: Nja jag tror vi har fått väldigt bra förståelse för hur det fungerar men vi försöker väl lite mer
198 komma åt den här beslutsprocessen och jag tänker väl lite så här... Jag antar att du får presentera väldigt
199 mycket data själv som de inte har brett om hur ser den processen ut och hur tänker du när du väljer vilken
200 data du ska presentera?**

201 Ja det får man välja efter man har suttit och pratat med dem och börjat förstå vad det handlar om. Jag vet inte hur jag
202 tänker, men man får vissa... är det ekonomi ja då är det en sak de är intresserade av och ja då är det vinsten och
203 marginalen, kostnad. Och sen får man sätta sig in i området helt enkelt.

204

205 **34:20 V: Hur arbetar ni internt med data, alltså inom ert företag? Analyserar ni den också eller?**

206 Nja vi har precis börjat med att analysera det ekonomiska och de olika sajterna och så. Annars är det väldigt dåligt, vi
207 har inte levt som vi har lärt.

208

209 **34:35 M: Känner du att ni inom företaget skulle kunna få många förbättringar genom att börja arbeta på det
210 här sättet?**

211 Vi sitter varje månad och klipper och klistrar i Excel och får hur mycket tid var och en har lagt på varje kund. Sånt
212 där finns ju färdigt så det är bara ett högerklick och exportera ut det om man har haft det. Men ja man hade kunnat
213 spara pengar absolut. Sen om det... Jag vet inte, vi kanske inte är, inte några stora grejer.

214

215 **37:10 M: Har du någon erfarenhet av stora mängder data kopplade till just marknadsföring?**

216 Du menar alltså en kampanj och så kollar man hur mycket man har fått sålt?

217

218 **M: Ja eller till exempel analysera vilka kunder man har och utifrån det göra mer riktad marknadsföring.**

219 Nej det området har jag inte varit inne på. Men det här är lite kul, det handlar om en tillverkare som, de gör såklart,
220 dels har de fem stycken, ungefär fem olika produkter igång samtidigt som skulle gå ut och bli [produkter]. Och på de
221 här då så gör man tester, jättemånga tester, mellan 5 till 10000 tester för varje produkt. Det verktyget som de hade då
222 där testerna sparades det gav inte stöd för att se, var är vi i helhet utan de fick bara veta för hur det har gått för de
223 tester som har gjorts den senaste veckan.

224

225 M: Så det gick inte att få någon överblick, man kunde inte dra någon slutsats utifrån datan?

226 Nej precis så då fick de spara ner varje vecka i Excelfiler och så fick de bygga ihop det där och titta på de olika
227 Excelfilerna. Det var liksom omöjligt för dem att få en överblick utan det bara försvann för dem i det här svarta hålet
228 som hette någonting, typ quality center. Sen hade de ingen koll då, så de tyckte det här var fantastiskt när de fick ut
229 datan och den blev visualiserad. De kunde jämföra produkter, hur lång tid testerna tagit och så här. Nu är det
230 jättegammal data så den kan inte vara känslig längre. Men då minns jag att de sa att " vi kör mellan 5-10 000 tester
231 och detta här tester som går här och det här... här måste du ha gjort fel [intervjupersonens namn], så där ser det
232 faktiskt inte ut. Det finns inte över 100 000 tester körda på någon [produkt] någonsin." Jaha nej men då får jag väl gå
233 hem och kolla då, det är mycket möjligt att jag gjort fel någonstans absolut. Då får man vara ödmjuk med kunden för
234 de känner ju sin data. Men så gick jag igenom skriptet och tittade i databasen och SQL-foldern och nej, nej det ser ut
235 såhär. Så jag var tvungen att säga till dem att jag hittar verkligen inget fel i mitt arbete. "Jaha men då får vi kolla
236 vidare på det". Det visade sig att det var att alla de här testade i Kina. Sen lite senare fick jag reda på att de hade löst
237 frågan då och då var det att cheferna i Kina fick bonus för hur många körda testfall de hade och kunde visa. Så de
238 kopierade samma testfall varje vecka och körde dem igen. Man får det man frågar efter. Sånt händer ibland, att man
239 hittar konstigheter som ingen känner till för att det varit dolt i datan.

240

241 Det ni ser här är för varje produkt så ser de trenden och här ligger man, de här projekten är i princip färdiga, man
242 ligger på 99% pass-rate. Test-coverage då är något som växer upp efterhand ju längre projektet går då. Där ser man
243 att en stor del av funktionerna är testade. De har ett annat system där de lägger in buggar och problem så här ser man
244 då hur många buggar som är förknippade med ett projekt då, ett projekt och då ser man att det växer och växer och
245 växer. Det är ju en tråkig trend, man vill ju se att det droppar av. Samtidigt ser man då att i detta projektet då att det
246 finns 6400 testfall som är blockerade på grund av buggar men det finns bara 965 olika buggar så samma bugg
247 stoppar många testfall helt enkelt. Så då gjorde jag en funktion där de kunde gå in och se... För denna produkt då så
248 ser man att här är en bugg som stoppar 417 tester just nu. Då är det den buggen man ska sätta sig och lösa först,
249 därefter ska jag lösa denna buggen för den stoppar 273 tester. Vi kan vänta lite med den här buggen eftersom den
250 inte stoppar så många tester just nu.

251

252 44:42 M: Så man får en ganska bra överblick över vad som faktiskt behövs göras?

253 Ja exakt, här har du ju prioordningen då. Om det är riktigt många träffar så är det någon funktion som inte är
254 färdigutvecklad eller koden inte är levererad än. Man kan gå vidare här någonstans... Här kan vi gå in och se vad det
255 är för liten bugg. Här ser ni legacy feature, ja det var ju inte så lätt att tolka vad det var. Här kan man då se att efter
256 [produkten] startats om så säger han att batterinivån är för hög, eller att den är högre eller lägre än vad den var precis
257 innan den startade om då. Då kan man ju vända på det för att se vilka tester är det som denna stoppar. Ja det är bara
258 ett enda här då. Man kan vända på det också och se hur många påverkas av en bugg eller vilka buggar är det som
259 påverkar mitt projekt mest då. Så både dem som sitter på buggdatabasen och de som sitter som testansvariga för
260 projekt de har en styrka här ju. Om vi tar och rensar allt här. Vi kan välja vecka... så här har vi då de olika projekten,
261 de är inte sorterade i passrate utan de är sorterade i bokstavsordning. Då ser man det exakt hur långt de har kommit
262 då och man kan ju välja om vi tar det här, så trillar vi ner en nivå och ser (???)47:28 i den produkten, hur långt har
263 man kommit i varje del och här är några som har det här och då är den ju nästan färdig. Basic bordering up den är
264 inte lika färdig då, så då kan man drilla sig ner där och se vad som inte är färdigt. Så har de typiska namn på testerna.
265 De ser då vilket område de behöver fokusera bättre på i testningen då och då får vi även progressen. Så här får man
266 då... [visar på skärmen].

267

268 48:30 M: Jag tror att vi har fått väldigt mycket bra material här och behöver inte ta upp hela din dag.

269 Nej men det är en kul grej. Den var ju lite gammal den jag visade. Så här är någon som lite nyare och fräschare. Här
270 är återigen att du kan välja... Här uppe har du hur många pengar man har sålt till en kund och här borta har du
271 marginalen. Här har du då de två största kunderna eller tre. Här kan du se vilka produkter vi säljer till de här största
272 kunderna. Här har vi någon produkt som vi har riktigt mycket marginal på och någon som vi säljer i väldigt stor
273 volym för mycket pengar men har inte så himla stor marginal på den. Så här kan man då fundera på om vi kan öka
274 marginalen på den då kommer det rassa på bra i kassan. Så ser vi över månaden hur det rör sig, trenden. Då får man
275 då överblicken att totalt sett har försäljning ökat 64% och marginalen öka 4. Då kan du gå ner och analysera en kund,
276 en produkt och där kan du gå in och ta fram liksom vilka fakturor har skickats med en produkten till den kunden. Du
277 kan få fram datum och prick pengarna då, du kan gå liksom från bolagsnivå, hur det påverkar bolaget, ner på enskild
278 faktura och se... Och det är där styrkan ligger då att du navigerar lätt tvärs igenom datan och ser... Du har både
279 helheten och detaljerna liksom.

280
281 50:41 M: **När man ser det här så känns det väldigt viktigt att som beslutstagare faktiskt ha digital kompetens**
282 **och förstå hur man ska använda den här datan och hur man ska navigera och så vidare.**
283 Ja, GUI:t ska ju vara såpass lättförståeligt att man ska kunna sätta sig och testa runt och förstå, det ska vara
284 intuitivt. Det är extremt viktigt, om man behöver skriva en användarmanual så har man misslyckats.
285
286 51:12 V: **Sitter beslutstagarna själva och kollar eller är det någon som kommer och presenterar datan för**
287 **dem?**
288 Vi har ju flera möten där man visar hur man har tänkt att det ska se ut och så här fungerar det, den här
289 funktionaliteten har det. Sen beror det helt på hur användarna, hur de är vana. Är de jättevana så är det inga problem
290 då kan de Qlikview och då är det samma funktion över allt, har du lärt dig hur det funkar så är du med på det. Det är
291 jätteviktigt att utbilda, annars får du inte det här intresset, annars blir det en för stor tröskel. Men jag upplever att det
292 är en väldigt liten tröskel då med Qlikview, det kan inte bli fel, du förstör ju aldrig någonting.
293
294 52:24 M: **Jag känner att vi fått svar på våra frågor så vi vill tacka så jättemycket.**
295 Ja vad bra, ni får lycka till och ha det bra!

Appendix 4 – Lundalogik

Interview by: Phone

Position: Deputy CEO

Date: 2016-05-02

M = Mattias

V = Viktor

-
- **Start of interview**

1 V: Berätta lite om din arbetsroll

2 Ja den är väl lite flytande. Men jag tar hand om all strategisk produktutveckling och marknadsföring. I
3 princip så kan man säga att jag... Mitt jobb är att se till att vi har en produktportfölj och en
4 marknadsportfölj som fungerar och som maximerar vår försäljning. Jag ska framtidssäkra kommunikation
5 och produktutveckling.

6

7 V: Hur arbetar ni då med data både internt och externt?

8 Det är ett ganska brett begrepp så den är inte helt enkel att svara på. Big data är ju... Om det är någon
9 som har en definition på hur mycket data det ska vara för att vara big data så hade det varit enklare.
10 Men det finns det ju inte. Vi jobbar med enorma mängder data i vårt perspektiv men de enorma
11 mängderna data är ingenting jämfört med... Jag var och tittade på ett visualiseringsevent med EF och
12 Maxlab förra veckan och jag skulle säga att fyra bilder från Maxlab producerar mer data än vad vi gör på
13 en vecka. Så att allt är relativt. Men vi jobbar ju med stora datamängder ur perspektivet att vi jobbar
14 med alla Sveriges företag och deras kontaktpersoner, eller alla nordens företag kan man säga. Vi har tre
15 produkter, alla inom sälj och CRM. En av de här produkterna är en salesprodukt, abonnemangsform.
16 Man skulle kunna säga att det är affärsvärldens Spotify. Vi har alla företags beslutsfattare på plats från
17 start och sen så uppdaterar vi dem varje dygn. Vilket innebär att vi har ju så mycket data som man kan få
18 tag på om beslutsfattare i Sverige, eller Norden. Så det är ju ett stort dataset ur perspektivet att det inte
19 går att få tag på mer men det är inte så jättestort ur ett sammanhang. Där jobbar vi väldigt mycket med,
20 framförallt att se till att den här är uppdaterad, att den är korrekt. Vilket oftast, i vår värld i alla fall är det
21 största problemet. Att de datamängderna vi har inte stämmer. Känner ni till Hampus Jakobsson som
22 grundade Tat?

23

24 V/M: Nej

25 Han är serieentreprenör. Just nu driver han ett antal bolag varav ett heter Brisk och de tog egentligen
26 fram ett beslutsverktyg som man åkte ut och presenterade för stora företag i USA. Det första de
27 upptäckte när de kom ut, så sa de att ja det här vill vi ha, det är jättebra men vi kan ju inte lita på datan
28 som finns. Så ni får börja i den änden. Så då fick de börja med att göra ett nytt verktyg så att de fick in
29 rätt data. Så det här är ett jätteproblem i vår bransch att väldigt mycket av datan man har i sitt system
30 är gammal och dålig. Det är egentligen skälet till att vi byggde den här produkten som uppdaterar sig
31 automatiskt. Ungefär hälften av alla poster som finns i vår totala dataset för den här produkten ändras
32 minst en gång i halvåret.

33

**34 V: Men den är fortfarande inte korrekt då, även om den uppdateras automatiskt? Det finns
35 fortfarande en del brister?**

36 Allt är relativt. Hittar ni alla telefonnummer ni letar efter på Hitta eller alla bolagen på Eniro?

37

38 V: Nej.

39 De finns ju, men vi hittar dem inte. Alltså är den då korrekt eller inkomplett. Det är en definitionsfråga
40 då. Sen vet man då att varje gång en telefonkatalog trycks så är antagligen uppgifterna korrekta,
41 antagligen säger jag. Antagligen korrekta när vi trycker katalogen. men så byter man jobb eller så händer
42 det någonting eller så har man en privat mobil och en jobbmobil så man tappar de numren. Sen så har
43 den individen ett annat kontaktnät än vad den hade tidigare eller så lägger man till en ny e-postadress så
44 det här är ju i konstant rörelse. Vi tror ju extremt mycket på att en av framgångsfaktorerna för företag i
45 framtiden är att lägga lite tid på administration. Det finns ett antal studier att en
46 informationsmedarbetare lägger i snitt två timmar per dag på att leta efter information. Då kan man säga
47 att om den informationen då inte är korrekt. Det ligger väl i det här mönstret att det läggs så mycket tid
48 där eftersom informationen inte är korrekt. Du hittar ett telefonnummer, ringer det och så får du inte

49 tag på personen. Vi försöker effektivisera företag genom att ta bort momentet av att behöva rätta lika
50 mycket data. Du kan titta på Lundalogik, vi växer med kanske 40 medarbetare förra året och vi växer
51 med någonting i den nivån i år också. Det innebär att ett företag som inte prenumererar på stora
52 datamängder eller ser till att det här finns i systemet. För att ha koll på våra medarbetare så ska ju dem
53 hålla koll på 40 nya människor på Lundalogik. Sen så tar du det gånger 1,2 miljoner företag i Sverige så
54 blir det här ganska så komplext och omfattande.

55

56 **M: Men hur fungerar det att datan uppdateras automatiskt?**

57 Automatiskt är ju att ta i då. Men för kunden upplevs det lite automatiskt. Vi samarbetar med Bisnode
58 som arbetar med företagsdata. Och de tar in uppgifter som patentbevisningsverket, bolagsverket,
59 skatteverket, de skickar ut enkäter till kunder och har team som sitter och uppdaterar det här manuellt
60 för att det är ju i grund och botten de som uppdaterar datan medan vi har en modell där vi får tillgång till
61 all den här datan. Och sen ser vi till att varje natt så gör vi stora körningar och uppdaterar den data som
62 förändras så att kunderna ska slippa göra det.

63

64 **V: Men är det också... Är digital kompetens ett problem, när du säger att de lägger mycket tid på att
65 hitta datan? Är det bristande kompetens eller mer att folk inte uppdaterar?**

66 Det finns många parametrar på det och jag vågar inte säga vad som kommer från det ena eller det andra.
67 Men vi kan väl säga såhär, har ni privata bilder?

68

69 **M/V: Ja**

70 När ni letar efter en specifik bild hittar ni den?

71

72 **V: Det beror på hur väl sorterade de är**

73 Handlar det om IT-kompetens eller handlar det om att vi har så mycket information att det är svårt att
74 hitta den. Vi vet alla att om vi kategoriserar bilderna, taggar dem, så är de lättare att hitta. men det görs
75 inte, man hinner inte, man prioriterar annorlunda. Där kommer ju en del av (10:19???) att har du
76 datasets, att hålla det uppdaterat, hålla det fräscht och hålla det så informativt... eller hålla det på en
77 nivå så att det är sökbart, det är extremt viktigt. Sen så har du nästa steg som man kan nå. I och med att
78 vi har all nordisk data, så vet vi också att i Sverige, Norge, Danmark framförallt så skiljer sig datan. Så vi
79 har samma uppgifter men den är strukturerar på olika sätt. I Sverige får vi oftast in den som absoluta tal,
80 medans i Norge får vi den i intervaller.

81

82 **M: Uppstår det att problem då när de är ostrukturerade, alltså när de kommer i olika format?**

83 Nja, problem, men det måste ju hanteras. Vill man söka ut ett bolag som omsätter 51 miljoner då måste
84 du ha en mjukvara i vårt fall som kan hantera att antingen söka... Låt mig säga såhär, när vi har ett fält
85 och vill hitta ett bolag som har 51 miljoner i omsättning så måste vi ha en mjukvara som kan hantera att
86 både leta upp de som finns i intervaller 50-55miljoner och sen så måste vi dessutom vända på det och
87 titta på att alla som har ett absolut tal som är inom det intervallet ger samma svar som om det finns i ett
88 intervall, om ni är med på skillnaden. Det är klart att det är som med allt sånt här, varje gång du behöver
89 ha någon form utav handpåläggning på data, att du behöver strukturera det själv manuellt, så finns det
90 felkällor. Och felkällor i datamängder... om du importerar som vi har 2,3 miljoner poster i Sverige, lägger
91 du handpåläggning på det och det blir fel så blir det väldigt fel. Jag tror att det var Volvia som för några år
92 sedan gjorde ett jättestort utskick där man hade råkat flytta en cell, man flyttade en rad nedåt. Så
93 skickade man ut ett väldigt stort utskick till väldigt många av Sveriges bilägare där man hade ett
94 personnummer som inte stämde överens med ens adress. Man hade då flyttat cellerna ett hack. Så det
95 finns ju riskmoment i det här, så kan man säga, varje gång det är handpåläggning. Det kräver ju en
96 ganska ordentligt process för att datan ska vara kvalitetssäkrad. Och det är väl en av de stora

97 utmaningarna med big data egentligen. Du kan vända och vrida på datan på 1000 olika sätt och få olika
98 svar på samma data. Det kräver lite självrannsaking.

99

100 **V: Har ni något sätt att kvalitetssäkra datan ni använder? Eller är det bara intuition?**

101 Själva indatan kvalitetssäkrar [Företag], det är det de jobbar med. De jobbar bara med affärsdata. Sen
102 har vi ju ett antal olika maskintester för att se till att data, egentligen då när vi importerar den så har vi
103 ett antal parametrar som vi tittar på. Är den importerad på rätt sätt, rent maskinellt och sen så gör vi
104 stickprov och sen så gör vi en manuell handpåläggning där vi tittar, "verkar det här stämma?" Men det
105 slinker igenom fel ändå. Vi importerar historisk data till exempel om alla företag. Så att du kan se vad ett
106 företag har haft för omsättningstillväxt de senaste 5 åren eller hur deras värdebild har förändrats. Igår
107 hittade jag ett bolag som var startat 2015 men som hade omsättning mellan 1998 och 2000, vilket borde
108 vara fullständigt omöjligt. Och idag ägnade jag mig åt att ta reda på om det är en bugg hos oss, är det
109 felaktig data eller är det fel i indatan. Så är det ju. Har du då riktigt riktigt stora mängder data, så är det
110 bara maskinanalys som gäller, det finns liksom inget annat.

111

112 **M: Vad använder ni på Lundalogik ert egna system till mest? Hur skulle du säga att ni använder den
113 insamlade datan för att skapa värde? jag förstår att ni säljer vidare er CRM-lösning, men jag antar att
114 ni manipulerar och använder er egna insamlade data också.**

115 Det gör man väl på massor av olika sätt. Givetvis för att bearbeta våra egna kunder, för att lära oss mer
116 om våra egna kunder. En extremt vanlig, eller ovanlig fördel för att kunder faktiskt vet vem som är deras
117 kunder. Man har en känsla av att vi säljer till B2B företag och huvudsakligen i storstadsområden. Men när
118 man... Eftersom vi har all data när man då importerar sitt eget kundregister, så är det väldigt lätt att
119 jämföra det här med hela Sverige universellt. Då kan man se direkt, vart man har marknadsandelar och
120 faktiskt vart mina kunder finns och var det finns potential. Det har generellt sätt företag ganska dålig koll
121 på.

122

123 **M: Det är väl lite det vi vill komma åt också med vår uppsats.**

124 Man kan säga såhär. Stora företag har alltid haft råd att jobba med dataanalyser, ta reda på var deras
125 frekvens finns, vilken potential de har. Medans små företag antingen av hävd, har trott att det är för
126 dyrt, att de inte har tid eller att de inte riktigt bryr sig, man hinner inte och man behöver inte.

127

128 **M: Sen är väl folk, folk förstår väl inte riktigt vilken ROI de kommer få på att faktiskt analysera de här
129 stora datamängderna.**

130 Nej men sen är det väl... Du kan ta oss som ett exempel, vi omsätter ungefär 150 miljoner. När vi
131 omsatte 10,15,20,25 miljoner så funderade vi aldrig på om marknaden är stor nog för tillväxten vi önskar,
132 därför att den var stor. Men om du omsätter 150 miljoner och är marknadsledare, vilket vi är i Sverige,
133 då kanske man börjar titta på, hur mycket kommer vi kunna omsätta här? Det kan vi bara göra genom en
134 analys, titta på hur stor del av kunderna har vi? Någonstans är det inbyggt att ju längre och ju större du
135 jobbar med någonting, desto större behov av dataanalys har du. Du har också mycket större
136 förutsättningar för dataanalys.

137

138 **M: Hur marknadsför ni er på Lundalogik? Kör ni via några sociala medier också eller hur fungerar det?**

139 En helt annan fråga och en ganska stor fråga. Men vi marknadsför oss i alla kanaler vi kan och sen så har
140 vi vissa prioriterade kanaler. Idag så går alla nästan via sökordsoptimering och sökordsanonsering och
141 det kräver väldigt mycket content. Vi jobbar med att skapa relevant innehåll och det marknadsför vi i de
142 kanaler som står till buds. Sen om det är seminarier eller mässor där du når en liten publik, men ofta
143 ganska specifik. Så väldigt mycket är ju sociala medier, Google, alltså allting. Se till att allt är sökbart och
144 hamnar högst upp i träffresultaten är vad det handlar om. Och där är sociala medier... De rankas för

145 närvarande ganska högt av Google. Går du tillbaka 3 år så var det mycket viktigare att få inlänkar än att
146 ha relevant content. De algoritmerna, sök-algoritmerna undergår ju ständig förbättring eller förändring,
147 jag vet inte vilket vi ska fastna vid.

148

149 **V: Hur arbetar ni då med den datan ni får in där från...**

150 Vi analyserar alla in-flöden vi har, sen så tänker nog inte vi i termer av data när det gäller inflöden. Vi
151 tänker i termer av signaler, sen så är det ju data bakomliggande givetvis. Det vi försöker fånga upp är ju...
152 Vi försöker rikta olika kanaler för att se vilka signaler en kund sänder. Ta ett exempel, så fort någon
153 signar upp för att testa på en hemsida så har de här människorna cookies på sig, om man inte har spärr
154 på det. Då tittar vi ju bakåt på vilka sidor har man besökt hos oss, när gjorde man det, sen tittar vi på hur
155 länge man spenderade tid på respektive sida och sen så väver vi samman massor med olika faktorer. För
156 att se hur väl, hur hett ett sånt här lead är. Sen så nästa steg så jämför vi det här med den data som vi
157 har... Eftersom vi har all Sveriges data så jämför vi det här med vår profil, vår kundprofil. Så någon som
158 kommer in och besöker oss lämnar ett avtryck, det här har jag gjort på sidan, det matchar vi med, är det
159 här en kund eller är det här ett prospect som matchar våran kundbild. Och det vi jobbar med just nu är
160 machine learning där vi tittar på om vi dessutom kan prediktera utifrån de här mönstrena, hur sannolikt
161 det är att det här kommer bli en kund eller inte. Allt det här viktas vi samman till ett värde som vi säger:
162 såhär hett är det här leadet. Det är klart att allt det här data som vi analyserar, men vi väljer att kalla det
163 signaler.

164

165 **M: Varför väljer ni att kalla det signaler?**

166 Det är lite enklare för hjärnan att ta till sig. Det är ju en signal om ett lead på hemsidan faktiskt är
167 intressant för oss eller inte, att istället gå ut och prata med säljare om att vi har datapunkter som
168 indikerar att det här är ett potentiellt lead, då tycker de att va!? Om vi däremot säger att de här har
169 signalerat att de är heta. Så det är egentligen bara för att förenkla och förtydliga vad det handlar om. Vi
170 har ett tusental besökare varje dag på hemsidan. De allra flesta av dem är helt irrelevanta och har
171 hamnat på sidan av misstag. Man har kanske sökt på ett sökord, klickat på en länk, nej det var inte det
172 här jag letade efter så är man därifrån på 3 sekunder. Tidigt i begynnelsen... Backar du 5 år så fanns det
173 ett antal verktyg Habsbo, Prospect, Prospect finder och liknande som då trackade trafik på en hemsida.
174 Sen så klär dem IP-nummer på det här och så jämför de IP-numret med företagsdatan och säger att nu
175 har Lundalogik besökt den här sidan. I början fick man långa listor på alla besökare, steget efter det var
176 att man fick listor där man hade sagt att de här sidorna är viktigare för oss som en kund besöker. Sen så
177 satte man ett värde på det, sen efter ett tag så fick man listor på värdet på besökarna. Nu så jobbar den
178 typen leverantörer på att titta om det här är relevanta besök eller inte relevanta besök, är de
179 återkommande besökare? Så att de själva jobbar med intelligens. De börjar inte mäta förrän du har
180 besökt två hemsidor eller två sidor på din hemsida. Det vill säga att du måste ha gjort ett aktivt val att
181 läsa mer, annars är det inte ett relevant besök. Det vi gör är att vi egentligen adderar och tittar på att
182 okej, passar det här besöket våran kundprofil? Sen så tittar vi, ja det passar vår kundprofil och sen så kör
183 man machine learning ovanpå det. då kan titta på att dessutom har de en tillväxttakt, lönsamhet och
184 många andra parametrar som indikerar att de här är ett väldigt hett lead. Man lägger hela tiden på mer
185 data för att förstå mer vad det är som händer. Men för en säljare som bara vill sälja så är det ganska
186 komplext, de bryr sig egentligen inte överhuvudtaget vad som har hänt, men om jag säger att det här är
187 ett hett lead så ringer de.

188

189 **M: Vi pratar mycket om att det är stor bristande digital kompetens bland de här större beslutstagarna i**
190 **alla svenska stora företag. Men som det låter på er, det låter som att ni ligger väldigt långt fram i tiden**
191 **just med den här datainsamlingen och så vidare. Men vad tror du att det krävs för en beslutstagare i**

192 **form av digital kompetens för att faktiskt förstå vad man kan dra för nytta av att samla in den här**
193 **datan och så vidare? Och att faktiskt ställa rätt frågor?**

194 Det är en ganska stor fråga. Jag kan säga såhär. Jag tror att det krävs ett extremt öppet sinne och en
195 förståelse för att om man inte har rätt verktyg i verktygslådan så får man se till att skaffa sig den
196 personalen och lita på den personalen som gör de valen. Vi kan ju se, tittar vi bara de senaste 5-6 åren så
197 har ju vår kravprofil på en medarbetare på marknadsavdelningen förändrats radikalt. Nu anställer vi
198 ju, datavetare, ingenjörer, systemvetare till vår marknadsavdelning där vi tidigare anställde människor
199 som var kreativa och duktiga på att skriva text. Nu har vi fortfarande kvar de som är kreativa och duktiga
200 på att skriva text, men de kompletteras med dataanalytiker. Nu har jag en kille som jobbar heltid med att
201 bara analysera mönster, data. Eftersom väldigt stor del av vår marknadsföring går ut på att skapa
202 relevant content så måste vi se till att det här contentet är sant. Vi använder väldigt mycket av den datan
203 som vi får in till att faktiskt ta reda på olika mönster och ska man då sitta i en ledningsgrupp och bevilja
204 att man anställer en heltidsanalytiker och man kanske anställer en civilingenjör till marknadsavdelningen.
205 Då måste man någonstans få bilden att det här är relevant, det finns pengar in här någonstans. De allra
206 flesta ledningsgrupper är väldigt ekonomiskt krassa och tittat man på det lite äldre gardet så kanske det
207 är så att man ännu mera säger att vi har gjort såhär, det fungerar. Det man behöver göra idag är inte att
208 utbilda alla inom IT och få dem att förstå det här utan det är att få alla att öppna sinnet och säga att det
209 finns nya sätt att jobba. Det tycker jag nog... Nu kanske vår kundkrets är lite speciell för de kommer till
210 oss för att de har upplevt ett behov, alternativt för att de har läser vårt innehåll oh tycker att det är
211 relevant att klicka på. Så vi får en ganska snedvriden bild. Någonstans så verkar ju alla som kommer till
212 oss intresserade av det här området. Det behöver inte vara så, vi har fem och ett halvt tusen kunder. Det
213 innebär ju att vi har 1,2 miljoner företag som inte är kunder. Det är mycket möjligt att de inte alls
214 motsvarar den bilden som vi har, så att gissningsvis är det så att det finns en bristande IT-kompetens och
215 den måste kompletteras med ett öppet sinne. Att okej det händer något här, vem kan hjälpa mig?

216

217 **M: Som ett exempel så pratade vi med ett lite större svenskt reklambolag förra veckan. Där vi då fick**
218 **pratade med en strategisk planerare och då pratade vi om big data och vi trodde att de arbetade mer**
219 **utifrån datainsamling osv. Men som det lät på honom så låg de långt efter på den punkten. Han ansåg**
220 **ju inte heller att de kommer att börja jobba med det mer i framtiden utan han såg et lite mer som en**
221 **fluga.**

222 Som med alla trender så har de en tendens att hypas och sen så går de tillbaka och blir mer normala. Så
223 svänger det här fram och tillbaka, pendeln svänger hela tiden. Big data är inte riktigt lika hett idag som
224 det var för tre år sedan, som term. Det är av den enkla anledningen att nu börjar folk ta till sig det här
225 och det normaliseras, det är en del av vardagen. Vissa trender vissa hypar hoppar vi också över. Vi var
226 jättetidiga på något som hette VAP en gång i tiden, som var föregångaren till 4g, 3g och 2g, långt innan
227 det och då kunde man skicka datatrafik. Där var vi tidigt på, det var jättebra och vi sålde en sådan
228 lösning. Men det var superhypat då, alla pratade VAPoch nu pratar alla Big Data. Oavsett vilken sådan
229 här trend eller hypertrend det tar så... de är ju ett steg i utvecklingen och man kan inte ta alla steg utan
230 ibland måste man ta två steg. Men om man bortser från datamängden så tror jag... Jag tror att det är
231 enklare att kolla på det här att man måste vara metrics driven mycket mer idag. Det är svårare att gå på
232 magkänsla, det är svårare att räkna hem saker baserade på någon form av magkänsla. Allt vårt
233 försäljningsarbete, allt vårt marknadsarbete, allt vårt analysarbete och utvecklingsarbete grundar sig i
234 dataanalys och det måste gå att räkna hem. Det tror jag många företag kommer få lära sig den hårda
235 vägen att det här måste man göra. Sen så kommer säkert riktigt kreativa bolag, alltså om du går till en
236 byrå som bara har i uppgift att vara en kreativ byrå. Då behöver du känna till de här termerna men du
237 kan inte typ ignorera dem därför att de inte är ditt expertisområde. Det är klart att det finns väldigt
238 många människor som klarar sig utan den teknik som finns idag och det går alldeles utmärkt eftersom de
239 har helt andra förutsättningar. Jag tror väl kanske att Big data är någonstans en förutsättning för att du

240 ska kunna jobba med AI och machine learning men effekten av AI och machine learning kommer att ha
241 mycket större impact på människors vardag än vad big data har haft. Men med det sagt för att det är det
242 folk kommer känna och se för att det är då det hamnar i händerna på dem. Nu vet jag inte hur många
243 som använder Siri, men Siri bygger ju på att det finns stora datamängder tillgängliga som hon... Det är
244 helt sjukt när man börjar prata om henne som hon också. Men som den här AI:n den artificiella
245 intelligensen som nyttjar den här datamängden, det tror jag inte att den vanliga konsumenten funderar
246 kring, att det är big data som möjliggör det. Man kan välja att ignorera det och sen så kan man få åka
247 med på resan eller så är man tidigt på det. Jag tror att vi kommer att hitta mer framgångsrika företag
248 som förstår det här och jobbar med det aktivt.

249

250 **V: Hur har då dina arbetsuppgifter förändrats sedan ni började arbeta med big data?**

251 Om jag bara vill kunna hänföra förändringen till big data så hade det varit enkelt, men om vi nu säger att
252 vi gör det. Vi har gått ifrån mindre punktinsatser, mindre mässor och mindre seminarier. Utveckling som
253 känns bra där man har rundabordsdiskussioner och sitter ner med ett fåtal människor till att man idag
254 tittar på... Vi samlar ju in användardata om allt användande av produkten och så tittar vi på vilka
255 funktioner som används, vilka som inte används. När vi bygger hemsidor så bygger vi ofta tre parallella
256 hemsidor som vi abc testar så ser vi vilket som vi har bäst utfall på. Så jobbar vi bara med den som har
257 bäst utfall och sen så itererar vi vidare på den. Mitt jobb har gått från att vara väldigt mycket känslöstyr
258 till att vara helt och hållet datastyrt. Sen så får man utlopp för den kreativa sidan genom att man får
259 skapa bra innehåll och bra content. Jag skrev ju... Nu bloggar jag i princip ett par gånger varje vecka eller
260 skriver debattartiklar varje vecka, det gjorde vi kanske två gånger varje år tidigare. Så att det har
261 förändrats extremt mycket. Men det krävs samma kunskaper, men du använder dem på ett annat sätt.

262

263 **V: Det sätter också mer krav på att analysera data rätt också.**

264 Gör man fel dataanalys och skriver en debattartikel i dagens nyheter så skulle det inte vara roligt.

265

266 **M: Vad ser du för koppling mellan just företagskultur och insamling av big data?**

267 Jag vet inte om det har med kulturen, men det har det väl indirekt. Om du som ledningsgrupp måste ha
268 ett öppet sinne till nyheter så måste du väl ha en kultur som är bejakande, du kan aldrig sitta still. Du kan
269 aldrig ha en företagskultur som säger att det där är inte för oss, det är okej att strunta i det. Då har man
270 nog lite problem. En företagskultur som bygger på nyfikenhet. Big data är bara ett steg, nästa år kommer
271 vi prata om något annat.

272 **Appendix 5 – Avensia**

273 **Interview by:** Phone

274 **Position:** CEO

275 **Date:** 2016-04-20

276

277 **M** = Mattias

278 **V** = Viktor

279

280

281 - **Start of interview**

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307 **M: Om du berättar lite först om din arbetsroll?**

308 Just det, jag är VD på Avensia som är ett av Nordens största företag när det gäller att leverera E-handelslösningar till
309 medelstora och stora företag samt stora handlare och B2B-företag. Det är väl min roll kort och gott.
310

311 **M: Hur arbetar ni då med datainsamling? Vad är det för data ni samlar in om era kunder till exempel?**

312 För vår egen skull så samlar vi in ganska lite data om våra kunder eftersom vi levererar E-handelslösningar. Däremot
313 så använder vi de enorma mängder data som våra kunder har samlat in om deras kunder som de i sin tur återanvänder
314 för att gynna deras kunder. För vår egen skull samlar vi in ganska lite data om våra kunder.
315

316 **M: Okey, vad är det kunderna ber er samla in för typ av data?**

317 Det är försäljningsstatistik, orderhistorik, kundbeteende, kundstatistik, segmentstatistik, logistikorderflöde etc. Så det
318 är rätt mycket statistik kring deras försäljning, vi bygger deras e-handels sajter och deras lösningar.
319

320 **M: Just det, så då använder ni i princip kundens sparade data och drar slutsatser utifrån det?**

321 Just det, för kundens egen räkning. Det är där vi främst kommer åt större datamängder och beteendedata och
322 liknande.
323

324 **M: Det är ju väldigt intressant, formar ni hemsidorna utifrån den data som ni får insamlad?**

325 Japp, det gör man dels helautomatiskt och dels semi-automatiskt man samlar information hela tiden om
326 kundbeteende så av alla miljoner kunder som har vart inne på vår hemsida så spårar vi hur de har klickat sig fram på
327 hemsidan, vad de har visat intresse för, vad de har lagt i kundkorgen, vad de har köpt, vad de har klickat på etc. Sen
328 använder vi det egentligen för varje sidvisning så tittar vi på vem du är som individ, hur du jämförs med alla som
329 någonsin har vart inne och sorterar därför informationen som ska vara relevant för dig. Så om du har köpt en iPhone
330 förra veckan så kommer du att se väldigt mycket iPhone-tillbehör på sajten i blickfånget. Om du visar dig
331 ointresserad av den här informationen så kommer vi börja visa andra grejer i stället. Det är inte lönt att tjata utan man
332 anpassar hela tiden. Det finns en realtidshantering av detta när det gäller kundbeteende för att få ut maximalt värde.
333 Sen har du ju en, den är helautomatiserad så den har ju med automatiska system att göra. Sen har vi en
334 semiautomatisk del där vi egentligen samlar ihop informationen och gör en mer kvalitativ av kunderna, vilket
335 segment de är med i osv. Sen beroende på den typen av statistik så kan du göra olika typer av utskick, riktade
336 kampanjer och den typen av actions inifrån och ut så att säga. Då är man inne på det halvautomatiska, då har du en
337 automatisk insamling men en manuell analys och halvautomatiska steg för att liksom återkoppla informationen.
338

339 **V: Hur anonym eller är det någon risk för den personliga integriteten?**

340 Man har eller kunderna är anonymiserade kan man säga, så man känner inte till speciellt mycket om kunden annat än
341 att den har en cookie och den här cookien har betett sig så här tidigare och den jämför sig så här med andra cookies
342 och därför ger vi den här informationen till den som har den här cookien. Tar du bort den här cookien så är du
343 anonym igen, då är du bortglömd då får du börja på nytt igen. Du kan naturligtvis välja att surfa i anonymt läge och
344 liknande så att ingen sparas osv. Om man tittar på de här klassiska PUL, personuppgiftslagsfrågorna som
345 personnummer och till och med namn osv är det långt ifrån att man alltid har eller ens använder det man har. Just
346 personnummer används extremt restriktivt, förutom kundnummer och liknande som uppstår inom exklusiva
347 kundklubber osv och där kunden själv har angivit det.
348

349 **M: Det låter på dig som att det är en win-win situation att samla in så här pass mycket data om kunderna, att
350 det gynnar såväl företag som kund.**

351 Ja, jag tycker nog att det är en win-win. Som kund vill man ha relevanta erbjudanden och information och snabbt
352 hitta rätt saker på sajterna så jag hade blivit rätt irriterad om sajter glömde mig helt och hållet. Som handlare är det ju
353 då en enorm fördel att, du har inte mer än 6-7 visningar att intressera en kund. Har du inte visat rätt saker så kommer
354 kunden att gå någon annanstans. Så som jag ser det så är det en win-win men det kan givetvis dras till extrema fall då
355 det inte längre är det.
356

357 **M: Så är det absolut, hur marknadsför Avensia sig?**

358 Vi marknadsför oss till andra företag då som säljer produkter online och vi har genom events och kampanjer och
359 liknande. Vi har ganska liten marknadsföring online och huvudsakliga marknadsföringskanalen är deltagande vid
360 events och liknande. Vi har ett eget b2bonline.se event till exempel som vi arrangerar själva och så deltar vi andra.
361 Tittar man online så är det nyhetsbrev på vår egen hemsida men det vi säljer är ju projekt i mångmiljonsklass så det
362 är mycket relationsförsäljning. Det är inget man köper av oss online.

363

364 M: Förlåt jag försöker hitta här när ni startade, när grundades Avensia?

365 Vi grundades, eller det beror på lite hur man räknar eftersom vi är ett samgående av tre bolag. Det första av de
366 bolagen grundades 1998, det första bolaget som hette Avensia startades 2001, så från 98 till 2001 beroende på vilka
367 rötter man vill spåra. Sen gick de här tre bolagen samman 2007.

368

369 M: Och ni har alltid sysslat med e-commerce så att säga?

370 De första åren så var e-commercen en liten del av erbjudandet så att säga. Då höll vi på med rätt bred IT-konsulting,
371 sen sedan 2011 så jobbar vi bara med det. 2015 sålde vi dessutom ett par dotterbolag som hade annan verksamheter
372 så det har varit en gradbit renodling men vi har hållit på med ehandel sen 2003 i någon slags skala.

373

**374 M: Det jag vill komma åt lite är väl egentligen alltså hur ert arbete har förändrats i och med den här big data
375 revolutionen så att säga, hur det har blivit mer möjligt för företag att samla in större mängder data om sina
376 kunder?**

377 Där har det ändrats. Hos oss har det vart en differentierande faktor och vi var rätt så tidiga med detta med
378 personalisering och insamling av kundbeteende så vi låg ganska före the crowd 2010-2011. Då var det nästan ingen
379 annan som gjorde detta och vi gjorde det bra. Så för oss har det varit en stor affärsdrivare. När det gäller vad vi gör
380 för kunderna så ändras det ju precis hela tiden. Just på big data sidan så har det inte ändrats så mycket på de senaste
381 åren utan vi samlar in rätt mycket information, sen själva den här semiautomatiska analysen och manuella analysen
382 som jag pratar om gäller kunderna främst. Vi visar dem hur man gör, vi visar dem hur man kan göra men det är sen
383 när vi väl säljer den tjänsten så att säga.

384

**385 V: Har det då vart något från kundernas sida någon skepsis när vi väl började presentera det här, att vi kan
386 samla in det här och det här om era kunder?**

387 Ja, skepsisen kom väl i början av våra kunder, uppvisar kunderna såna mönster eller kan vi då dra några giltiga
388 slutsatser från detta? Ska man verkligen segmentera på en användares nivå? Det är egentligen en klassisk
389 segmentering har man gjort ett påstående och sen sökt folk. När man liksom då tittar på det här, damer i
390 femtioårsåldern handlar annorlunda än herrar i trettioårsåldern. Det har man hållit på med ganska länge, däremot att
391 som vi gör nu att vi egentligen går ner på individen, att du som har det här beteendet är sannerligen intresserad av
392 andra som har samma ungefär samma beteende. Den är ju rätt annorlunda och den kräver rätt stor datamängd så den
393 fungerar ju inte heller hos riktigt alla våra kunder alltså eftersom de inte samlar in tillräckligt med data. Så för dem är
394 det smartare att använda mer semi-manuella eller helt manuella metoder så det har vart en resa där att både övertyga
395 en del att det här fungerar, att det är värt att göra och sen också bevisa det helt enkelt.

396

397 V: Är det problem då med den digitala kompetensen hos kunder?

398 Jag skulle säga att när det gäller big data så och om man ska titta på den automatiska sidan så fungerar det och det
399 har fungerat i ganska många år för att det är där, där behövs inte jättemycket kompetens, man behöver skicka
400 systemet egentligen för att det ska framföra en egen verksamhet. Så att det finns mycket kompetens externt att hämta
401 och sen när det väl är automatiserat så kör det egentligen utan kundernas manuella inblandning. Det semiautomatiska
402 däremot och riktiga big data analyser med data mining osv, där skulle jag säga att det går extremt långsamt. Det är
403 personalkrävande, det är kompetenskrävande och jag skulle säga att i alla fall i dagsläget så används kanske högst
404 1% av den data som samlas in används faktiskt och om man skulle liksom ta det en division till och tänka på hur
405 många slutsatser man skulle kunna dra så är det enormt underutnyttjat. Det samlas väldigt mycket data, det används
406 väldigt lite data.

407

408 V: Vet folk vad de vill ha för data?

409 Leverantörer har sålt big data länge, framför allt IBM och den typen av databasföretag, Oracle och de andra stora
410 drakarna. Jag skulle säga att kunderna har fortfarande en rätt lång väg på den manuella sidan att komma ikapp.

411

412 V: Men vet de vilken data de vill samla in, eller samlas det in allteftersom?

413 Det är egentligen, där slår du huvudet på spiken egentligen. Det är, du samlar in för att veta exakt vilken data du vill
414 samla in så bör du kunna din verksamhet utan och innan. Det är lärande på något sätt. När du börjat samla in börjar
415 du ana vad det är du vill samla in. Sen får jobba vidare iterativt och borra och sen får man inte glömma att slänga lite
416 också och sluta samla in en del.

417

418 **M: Känner ni då från Avensia att när ni utformar sidorna, har den kompetens som krävs för att utnyttja den**
419 **datan som ni får tillgång till?**

420 Jag tycker att mycket av det är verksamhetsstatistik så det innebär att många saker som är gemensamma för många
421 verksamheter har vi kompetens kring. Andra saker där när det är branschspecifika beteenden där det är väldigt
422 nischade saker där vet vi ju att vi får lära oss tillsammans med kunderna. Stora saker och stora gemensamma
423 beteenden och trender känner vi ju till. De stora sakerna vi har jobbat med hos kunderna känner vi ju då också till.
424 Men det kommer hela tiden nya saker och man ger sig in i nya branscher och hamnar i nya frågeställningar och får se
425 saker man aldrig sett förut, där får man vara ödmjuk och inse att man kan inte allt och får lära oss efter hand. Men
426 det är klart, det bygger upp en erfarenhetsbank.

427
428 **M: Jag tänker nu när du driver Avensia framåt, hur gör du någon omvärldsanalys? Vilken data samlar du in**
429 **i det sammanhanget?**

430 Från Avensia så är det viktigare att kolla på dels produktutvecklingen, vilka produkter och erbjudanden finns det där
431 ute? Vi kollar också på konkurrensen, vad erbjuder våra konkurrenter, vilka är våra konkurrenter? Och sen även
432 kunderna, vad efterfrågar kunderna? Vad ser vi för trender på kundsidan? Det här är egentligen rätt så kvalitativ
433 kartläggning, man deltar på events, mässor och jobbar med sina nätverk. Det är inte så mycket big data från vår sida.
434 Vi säljer ju tjänster till ett tjugotal kunder osv så för oss är vår datainsamling kring grunderna traditionellt CRM
435 skulle jag säga. Att kalla det big data är lite överdrivet, det är ganska enkelt.

436
437 **M: Men tror du att det finns bra potential för en big data lösning för Avensia? Tror du att det skulle leda till**
438 **att du som VD tar bättre beslut?**

439 Det som skulle vara intressant för oss skulle vara att skapa någon slags gemensam data över alla våra kunder och
440 deras kunder så att säga. Då har man ju vissa rättighetsproblem kring det. Men det finns ju produkter vi använder
441 som skulle kunna samla aggregerad information över kundgränserna. Där skulle man ju framför allt kunna hjälpa
442 våra kunder att maximera sin försäljning ännu mer. Problemet är ju då att användarna, slutanvändarna har sällan
443 samtyckt till att information får användas mellan olika handlare. Vilket innebär att du direkt kommer trampa in på
444 legala frågor och i vissa länder är det explicit förbjudet att göra sådana saker. Men där finns en enorm potential om
445 vad man skulle kunna samla för kundinformation och beteendeförändring mellan kunderna och det skulle verkligen
446 kunna utveckla Avensia framåt. Då pratar jag inte om mina kunder, utan då pratar jag om mina kunders kunder. Det
447 är där vi skulle kunna lära oss väldigt mycket.

448
449 **V: Är det så att många av era kunder är direkta konkurrenter med varandra?**

450 En del är ju det, samtidigt så skulle man ju kunna dra lärdomar över branschgränser också. En del lärdomar bär inte
451 över branschgränserna men en del gör det. Då skulle man kunna göra en kvalitativ granskning där också. Vilken ska
452 vi ens satsa på eller vilka saker som går över branschgränser men det skulle vi definitivt kunna lära oss, det skulle
453 kunna vara ett enormt värde för oss att kunna bygga upp.

454
455 **M: Jag tänker sen hela den här big data eran har kommit, anser ni att ni har fått nöjdare kunder?**

456 Jag vet inte riktigt, man flyttar ju fokus. Man lägger hela tiden till nya funktioner och sajterna blir bättre och mer
457 intelligenta så att säga men det blir deras konkurrenters sajter också. Det är nog svårt att säga att våra kunder har
458 blivit mer nöjda men kundernas kunder däremot, slutkonsumenter har blivit betydligt mer nöjda idag med
459 webbsajterna än vad de var för några år sedan och detta till stor del på grund av big data och beteendeförändring.

460
461 **V: Hur ser beslutsprocessen för dig ut? Finns det något specifikt kring hur du går tillväga när du fattar något**
462 **större beslut.**

463 Jag kan inte påstå att jag använder big data, jag utgår mer ifrån klassiska nyckeltal och redovisningsinformation helt
464 enkelt. Jag har inte som sagt, mina kunder, de stora ehandlarna. Jag har inte jättemycket statistik om hur dem ringer
465 in till oss eller hur de föredrar att bli behandlade. Vi har kundnöjdhetsundersökningar men det är allt, det är bara en
466 sammanfattning så det är inga stora grejer med det.

467
468 **V: Hur mäter ni kundnöjdheten?**

469 Vi använder 18:38??? Finns på nätet, rätt standardiserad och analyserat sätt att mäta.

470
471 **V: Vad ser du för möjligheter med big data i framtiden?**

472 Det finns väl alla möjliga möjligheter, alltså man har ju rätt länge sett ehandeln som en postorder i princip. Men nu är
473 ehandeln en ganska vass säljare och imorgon kommer ehandeln vara med all säkerhet den absolut bästa säljaren alla

474 kategorier som kan bemöta kunden precis som kunden vill bli bemött. Som vet allt som kunden, kan lägga fram
475 relevanta förslag och naturligtvis så undviker man att kunden ska känna sig integritetskränkt osv. Jag tror att big data
476 både realtidsbehandlad och semiautomatiskt behandlad kommer att lyfta fram ehandeln långt förbi vad en manuell
477 säljare kan, alltså vad en mänsklig säljare kan prestera.

Appendix 6 – King

Interview by: Phone

Position: Planner

Date: 2016-04-25

M = Mattias

V = Viktor

-
- **Start of interview**

1 V: Kan du berätta lite om din arbetsroll?

2 Absolut, jag är planner på reklambyrån King och vad det innebär i breda lag så innebär det att det är jag och våran
3 avdelning som sköter marknadsstrategierna sätter riktningen för det kreativa. Det innebär väl egentligen att om en ny
4 kund kommer till oss och säger att vi vill göra kommunikation så är det vi som försöker filtrera vilka utmaningar,
5 problem och därmed möjligheter som finns och så man bör rent kommunikativt exploatera. Så att all kommunikation
6 på något sätt har någon affärsrelevans så att vi liksom gör konton som gör skillnad istället för att bara göra roliga
7 eller bra idéer utan ett djupare syfte eller mening.

8

9 V: Hur avgör ni liksom vad som är affärsrelevant då?

10 Nej men det är väl... Steg ett är väl någonstans att... för vår kund att säga att de har ett problem eller att de har en
11 utmaning eller snarare säger dem att de vill göra kommunikation. Och då så säger vi okej ni vill ha kommunikation,
12 givet eran produkt så ser eran marknad ut på det här sättet, det här är era konkurrenter och de fungerar och agerar på
13 det här sättet. Det här är era konkurrenter och de fungerar och agerar på det här sättet. Det här är era konsumenter de
14 fungerar och agerar på det sättet, de tycker det här om erat varumärke, det finns de här trender eller strömningar i
15 samhället och så vidare. Massa olika parametrar. Utifrån det så får vi någon form av nulägesbild av varumärket eller
16 produktens roll på marknaden och vad man har för utmaningar och problem framöver. Om utmaningen är att
17 konsumenten inte gillar varumärket för den här och den här faktorn. Eller är utmaningen att priset gör att man inte
18 värderar produkten som andra substitut eller alternativ. och så vidare. Det är en uppsjö av lika problem men lika så
19 möjligheter. Konsumenter kvinnor i ålder 40+, av någon anledning förbrukar den här produkten mycket mer än alla
20 andra men ni har aldrig gjort någon riktad kommunikation gentemot dem. Varför inte utnyttja den möjligheten osv.
21 Med allt det här identifierat, med den analysen gjord så skapar vi ett internt verktyg här hos oss som även kunden tar
22 del av som vi kallar BPS (brand positioning statement) som i korthet blir liksom ett ramverk för kreatörerna att hitta
23 på idéer inom. Då blir det som en kedja, givet att marknaden ser ut såhär och ert företag står och gör det här så har ni
24 de här och de här möjligheterna och utmaningarna och därför behöver vi hitta på idéer som talar om det här och det
25 här. Så kan man säga i korthet. Och sen så gäller i princip sak om det är på kampanjnivå. Om [Företag] som är våran
26 kund vill prata om svenskt kött så kanske vi gör en analys för att komma fram till att konkurrenterna pratar om
27 svenskt kött på det här sättet och därför finns det en möjlighet att prata på ett annat sätt eller konsumenten just nu har
28 tittat för mycket på cowspiracy så att de vill inte ens köpa kött. Då kanske vi ska prata om det på ett helt annat sätt,
29 vem vet.

30

**31 M: Bra förklaring. När ni gör den här marknadsanalysen och den här målgruppsanalysen. Vad gör ni för
32 insamling av data då?**

33 Jag förstår att det skulle komma med tanke på att det var big data ni pratar om. Då kan man väl säga så här att i
34 mångt och mycket oftast så har inte vi interna verktyg här för att ta fram data. Utan ofta ligger datan redan hos våra
35 kunder. Tar man [Företag] eller [Företag] så har de väldigt omfattande databaser med all möjlig typ av information
36 och data. ibland får vi ta del av vissa saker de ger oss och säga att det här skulle vi vilja utgå från. Eller så får vi
37 tillsammans med dem jobba på att se vilken typ av parametrar man skulle kan filtrera lika data på och på så sätt hitta
38 olika insikter. Så oftast kommer den datan vi använder från våra kunder och deras egna databaser och sen så i den
39 mån som det är möjligt, och det är inte allt för ofta, så hittar vi egna rapporter genom olika rapportverktyg eller lite
40 smått googlar på det och hittar publika rapporter. Men det är vanligare att vi använder kundens data. Men... Då blir
41 det någon följdfråga, hur använder vi den datan som vi får av dem. Man kan säga så här, vad gäller våra analyser så
42 grundas de ofta av den datan men den data ger ju bara en bild av hur det ser ut. Alltså den här segmenteringen köper
43 de här produkterna vid det och det tillfället, i dem och dem geografiska områdena och så vidare. Men den säger
44 aldrig riktigt varför det är så. Så våra egna kompletterande analytiska metoder och verktyg blir djupintervjuer,
45 fokusgrupper, fieldstudies, allting för att få en bild av varför det ser ut som det gör. Data är ovärderliga för att bilda
46 sig en uppfattning om hur det ser ut och även hypotesen varför men det är sällan den bekräftar exakt varför. Där gör
47 ju vi mycket arbete genom att helt enkelt i grund och botten prata med konsumenter och försöka hitta anledningar till
48 varför saker och ting ser ut som de gör.

49

**50 V: Brukar ni komma fram till saker som era kunder inte själva lyckas komma fram till? Gör ni en mer
51 djupgående analys än vad de har gjort innan?**

52 Både och, ibland handlar det om att ställa frågor på ett sätt som de inte brukar ställa frågor. Ofta är det så att de flesta
53 stora företag som jobbar intensivt med kommunikation har olika typer av tracking verktyg och tracking rapporter.
54 Alltså använder sig typ av (???:40) liknande och får där med rapporter som säger att liking är så här, brand
55 awareness är så här, avsändarkopplingen är så här och så vidare. Även de svarar inte alltid på varför det är så, snarare
56 svarar de väldigt sällan på varför det är så. Så därför hittar vi egna sätt att komma till svars som de kanske inte gör.

57 Frågor de kanske inte vill ha svar på, givetvis i samråd med våra kunder så att de vet vad som försigår. Men det
58 handlar också våran process där vi måste ha friheten att ställa frågor på ett sätt som kanske är annorlunda än vad de
59 skulle ställt frågan på. Sen ska man också säga att många av de rapporter som görs hos företagen, hos våra kunder är
60 i standard format, det är standardrapporter med, självklart tweakningar beroende på vilken bransch och vilket företag
61 och så vidare. Men det är någonting som görs kontinuerligt. När vi gör våra har vi ofta ett mer riktat syfte och en mer
62 tydlig frågeställning och därför kan vi också gå lite mer på djupet inom den. Långt svar, ja ofta kommer vi fram till
63 andra saker som de inte kommer fram till och vi ser kopplingar som de inte sett och kanske utmanar de slutsatser de
64 har dragit av en statistisk analys de har gjort och ibland är vi på samma sida från första början. Men då handlar det
65 fortfarande från vårt håll att få den bekräftad genom att då göra intervjuer och liknande. Var det ett svar på frågan?
66

67 **M: Ja. Till exempel med fallet [Företag]. Vad får ni tillgång till, får ni tillgång till till exempel kundtrender de
68 har hittat i deras analyser eller vad är det ni får tillgång till?**

69 Ja. Vi får väl... det är inte så att hela deras databas ligger öppen för oss. De är ofta ganska komplexa databaser och
70 man måste lära känna hur man använder dem och så vidare. Oftast får vi be om saker som vi undrar över men när de
71 själva kommer med egna slutsatser och analyser och trender osv. Så är de givetvis måna att dela med sig av dem till
72 oss precis som vi är måna att ta del av dem. Eftersom vi i slutändan vill göra kommunikation som gör faktiskt gör en
73 förändring och göra skillnad. Därför handlar det om att försöka ha en så öppen delning av information som möjligt.
74 Nej vi kommer inte åt varenda liten grej och vi vet nog inte varenda liten grej som de skulle kunna pruta ner på men i
75 mångt och mycket så har vi en dialog där vi kan segmentera och få ut ganska specifika svar.
76

77 **V: Hur ser processen ut, när ni får den här datan? Hur ser själva analysprocessen ut? Har ni en analytiker
78 som försöker dra slutsatser och sen skicka det vidare eller hur ser det ut?**

79 Nej det har vi inte. Vi är inte så stora att vi kan ha en hel, en speciell person som är avsatt till analysinformation utan
80 de är väl vi planners som sitter med det oftast och försöker bearbeta den data som vi får. Och där går vi tillbaka till
81 det här samarbetet för att oftast sitter våra kunder på desto större analytiska avdelningar så att de kan ju lära oss
82 väldigt mycket och hjälpa. Men sen så gäller det för oss att göra det hela till något som är applicerbart och
83 användbart i kommunikationen.
84

85 **M: Tror du att ni kommer arbeta mer dataorienterat i framtiden? För den trenden vi ser nu är att big data
86 växer mer och mer och att företag anpassar sig även mer och mer och skaffar sig nya datalösningar? Tror du
87 att King kommer att göra något liknande, att ni kommer ta del av mer information och kanske kunna se fler
88 kundtrender och så vidare?**

89 Ja det tror jag absolut. Med tiden så är det som du säger att big data är ju relativt nytt men även om man använt sig
90 av data länge så är själva begreppet big data relativt nytt och jag tror bara att det kommer bli ännu större. Men med
91 det så tror jag även att det kommer att komma verktyg som gör det lättare för fler att ta del av den kunskapen som
92 finns tillgänglig. Verktyg som gör det lättare att hitta den information som man skulle vilja använda sig av helt
93 enkelt. Därför tror jag att det blir lättare för oss att ta del av den också men det är ju den ena sidan av det. Sen så tror
94 jag rent kreativt, alltså våra leveranser, så tror jag att de också kommer bli mer... Göra mer nytta av den big data som
95 finns. Med lösningar som på något sätt bigger på stora data. Vi har ju [Företag] som kund och vi gjorde en kampanj
96 för dem som hette [Kampanjnamn]. Jag vet inte om ni sett eller tagit del av den?
97

98 **V: Nej...**

99 [Kampanjnamn] var egentligen så här att [Företag], sen flera år tillbaka har väldigt mycket data över hur deras förare
100 har kört, det vill säga GPS-data. Vart de plockar upp en produkt och vart de lämnar av en produkt och hur de kör dit
101 och så vidare. Hur mycket data som helst, som vi gjorde en idé av som var [kampanjnamn]. Vi gjorde en turistguide
102 utav den data som fanns. Vi kunde få ut data som visade att det här området i Stockholm är det område där flest
103 upphämtningar sker och det här området är där flest avlämningar sker och då kunde man utifrån det och utifrån
104 områden och utifrån de nöjen som finns i det området se vad som var troligt att folk gick till. Ett enkelt exempel,
105 Stureplan är säkert en jättevanlig avslutningspunkt i Stockholm och skulle man gå in då och så ser man att jag är ute
106 efter restauranger i det området och så kunde man zooma in och så... Ni får gå in och kolla så kommer ni förstå vad
107 jag menar. Men så kommer man då se att runt stureplan i Stockholm så är det mer intensivt med avlämningar av
108 kunder och vad finns det i det området. Då kunde man koppla samman den här datan med typ google maps eller
109 andra typer av kartor som visade vilka typer av grejer som fanns. Restauranger, biografier, nattklubbar, vad det nu
110 kunde tänkas vara. På så sätt gjorde vi en intuitiv turistguide över Stockholm som egentligen bara grundade sig på
111 vart stockholmarna själva åker. Så av datan vi fick av [Företag] som de har tillgängligt.
112

113 **M: Det känns som att uber har använt precis samma data för att optimera deras placeringar och deras bilar.**
114 **Det var väl jag tänkte på rent spontant.**

115 Ja men säkert, det borde de verkligen göra. De borde vara smarta nog för att se vart folk åker och vart folk hämtas
116 upp och sen försöka optimera sin verksamhet efter det. Men ett annat sätt som big data används på är ju inom
117 programmatic-marknadsföring. Alltså att finnas automatisk på nätet och på andra ställen bara baserat på vart ens olika
118 kunder rör sig helt enkelt. Jag skulle inte påstå att det är det mest kreativa sättet att göra kommunikation på men det
119 kan säkert vara väldigt effektivt beroende på vilken typ av information man vill sprida helt enkelt och vilken typ av
120 kunder man vill nå.

121

122 **M: Det är väldigt lätt att nå rätt målgrupp med den typen av marknadsföring.**

123 Exakt. Och den typen av lösningar blir kanske mer en mediebyrå som hanterar det. Eftersom det är de som bör veta
124 det. Vi är de som säger att här har ni en banner och mediabyrån säger att den borde vi lägga på de här och de här
125 sidorna. Eller digitalbyråer som hanterar adworks och så vidare. Det är intressantare för oss att vi ger ett... dels ett
126 insiktsperspektiv men också ur ett lösningsperspektiv hur man kan använda big data. Både de perspektiven kommer
127 att bli allt mer relevanta, för att inte säga nödvändiga för oss att ha en del av.

128

129 **M: Det är också en väldigt tydlig trend nu att många företag anställer just data analysts för att kunna se de**
130 **här trenderna och förstå hur de ska placera sig i framtiden, hur de ska utveckla sina produkter och så vidare.**
131 **Tror du att ni på King, skulle ni behöva anställa någon med det här yrket eller tror ni att ni har den**
132 **kompetens som krävs för att kunna ta nytta av big data?**

133 Jag tror inte att det är aktuellt för oss att anställa någon. Men jag tror absolut att det skulle kunna vara värdefullt för
134 oss att ha en sådan person, det tror jag. Men i den mån som vi behöver använda oss av data så tror jag att vi har dels
135 kompetensen själva att ta ut den data vi behöver och dra de slutsatserna som vi behöver dra. Har vi inte det så tror jag
136 framförallt att i den närmsta tiden fortfarande handlar om ett nära samarbete med våra kunder. Det kan ha att göra
137 med att vi är en mindre byrå sett till manskapet, en större byrå kanske säger annorlunda, det vet jag inte. Men jag
138 tror inte att det är riktigt nödvändigt för oss än.

139

140 **V: Du har alltså inte upplevt att ni haft något bristande kompetens? Alltså att nu hade vi egentligen behövt**
141 **den här kompetensen för att få reda på det här.**

142 Nej, det tycker jag faktiskt inte. Man kan ju hitta jättemycket intressanta insikter utifrån big data, alltså hur mycket
143 som helst. Men jag tror att man kan komma åt den genom våra kunder som sagt. Ett annat sätt att använda big data i
144 insiktssammanhang är ju att försöka på något sätt... Typexempel av vad man kan använda big data till är ju för att
145 hitta vilka som är ens lönsamma kunder. Ni känner ju till den här 20/80 regeln att det är ofta 20 % av kunderna som
146 står för 80% av omsättningen. En typisk sådan insikt behöver inte vara så svår att komma till utan det är ganska
147 simpel... inte simpel men det kan vara ganska tydligt och det kan vara ganska enkelt att se. Och det klarar vi av. I det
148 fallet blir våran roll mer fokus på att lyfta upp den frågan för våra kunder än att hitta den från första början. Så
149 kunderna kan vara såpass insiktsfulla att de själva inser det men samtidigt inte av bekvämlighetsskäl eller av
150 seghetsskäl eller vad det nu kan tänkas vara. Inte vågar lämna det skrå de är i och se till nya kunder.

151

152 **V: Använder ni någon typ av data för att rikta er till nya kunder? Försöker läsa av vilka företag som är**
153 **intressanta?**

154 Nej, det gör vi inte. Våra nya kundsarbeten uppstår oftast genom dialogkontakter, dialog och annars pitcher. Jag
155 tror att det är svårt att mäta och på något sätt kvantifiera kreativitet eller behov utav en viss typ av kreativitet utan det
156 handlar mer om någon form av matchning. En inställning till kommunikation och en samsyn på problemet. Jag tror
157 inte att data tyvärr skulle kunna hjälpa oss att komma till insikt med de grejerna.

158

159 **V: Hur har ditt arbete förändrats sen ni började arbeta med big data eller sedan det började bli en grej?**

160 Jag tror inte att den har förändrats så mycket, annat än att det finns mer tillgängligt och därför blir processerna mer
161 komplexa och mer intensiva. Det finns mer data tillgängligt och det finns mer insikt att ta lärdom av men då finns det
162 också mer beteenden och hypoteser att bekräfta. Så att den har hjälpt oss på så sätt att vi tidigare i processen kan bli
163 mer riktade mot det vi ska åstadkomma.

164

165 **V: Anser du att big data har hjälpt konsumenterna också? Att ni kan göra en bättre reklam som inte bara**
166 **gynnar era kunder utan även konsumenterna?**

167 Absolut. Men det beror ju på hur du definierar gynnar konsumenterna också. Men jo det skulle jag absolut säga för
168 att den blir mer relevant så utifrån vårt perspektiv så tycker vi det. Mer relevant för kunden för att vi har ett mer

169 relevant affärsproblem och den blir mer relevant för konsumenten är en viktig del i ekvationen av problemet. Men
170 också för att vi genom att lära känna konsumenten kan göra kommunikation som har smort(25:31???) bättre helt
171 enkelt. Så att ja, jag skulle säga det. Men samtidigt så kommer man in på en liten diskussion om hur pass relevant
172 kommunikation faktiskt är för konsumenterna.

173

174 **M: Men det är också en tydlig trend att det är en brist på digital kompetens hos de stora företagen. Känner**
175 **du att det finns något brist på digital kompetens hos King? Tror du att ni skulle gynnas av att lära er mer om**
176 **datainsamling och så vidare?**

177 När vi pratar digitalt så har vi en person som är liksom digital creative director eller liknande här. Så vad gäller
178 användandet av det digitala så har vi nog inget behov av det, vad gäller framtagningen och analysen av digital data så
179 jo absolut där känner jag att vi har mycket kvar att lära oss och fanns det någon som skulle kunna erbjuda den typen
180 av hjälp eller utbildningar så hade det varit intressant.

181

182 **M: Arbetar ni med sociala medier på något sätt?**

183 Ja. Inte för oss själva men för våra kunder så gör vi ju... Det ingår i vårt uppdrag att vi ska vara medieneutrala
184 ibland beroende på uppgiften och vad målbilden är så kan ju sociala medier vara en relevant kanal att använda sig av
185 och då hittar vi lösningar för det.

186

187 **M: Som jag har förstått det, när man väl skapar en kampanj på Facebook och så vidare, då får man välja**
188 **exakt vilken region, vilken målgrupp och alla de andra faktorerna som man vill rikta sin kampanj mot så att**
189 **säga. Då antar jag att ni utifrån den data ni fått in att ni bygger de resultaten på det?**

190 Ja. Alltså jag tror att vad gäller själva segmenteringen via facebook till exempel. Så har vi nog i ett mycket tidigare
191 bestämt att det här är målgruppen vi vill nå. Och det har vi bestämt utifrån data också men även utifrån en målbild
192 hos kunderna. Vad gäller själva genomförandet av segmenteringen när det kommer till att applicera det på Facebook
193 till exempel så är det ofta inte vi som sköter det utan en digitalbyrå, en mediabyrå eller de själva skulle jag säga. Men
194 sen så har vi ju diskuterat den tidigare så att vår lösning som sedan bara ska införas på Facebook är grundad på den
195 insikt och därmed segmentering.

196

197 **V: Anser ni att era kunder ofta har den digitala kompetensen som de behöver? Vet de vad för data som är**
198 **relevant och hur den ska hanteras?**

199 Ja. De har absolut relevant kompetens men sen så är det ju vårt jobb att komma ur ett kommunikativt perspektiv och
200 därmed också ur ett konsumentperspektiv eftersom det är de vi ska nå i slutändan. Det är vårt perspektiv på hur vi
201 ska göra saker. Sen kan vi komplettera dem genom att komma med nya typer av infallsvinklar eller utmana redan
202 framtagna resultat eller analysmetoder. Så ja, jag tror att de sitter på kompetensen men sen så kan det vara bra att
203 komplettera synsätten.

204

205 Jag tror att man ska ha med sig någonstans att insamlandet inte är gratis heller. Därför har vi svårt att samla in data
206 själva. Vi har vår egen konsumentpanel med 400 svarande som vi kan skicka ut undersökningar till och få data så att
207 säga. Men det är knappt nämnvärt jämfört med [Företag]s egen databas som är utifrån varenda gång någon drar sitt
208 kort i en butik. Vi använder oss av den data vi får av våra kunder eftersom den oftast är mycket större och sen får vi
209 använda våra egna metoder för att utveckla det ännu mer.

210

211 **M: Ja men precis.**

212 Men vad säger ni, alltså det här är bara en personlig åsikt. Men för mig så är det intressant men det har också blivit
213 en snedvriden diskussion om big data och hur stort det har blivit. För mig så tycker jag att företagen har alltid gjort
214 det men ju mer tekniken utvecklas desto mer data kan man samla in på ett enkelt sätt. Och sen så har man då också
215 insett att ju mer data man kan samla in desto mer insikter kan man få ut av det. Det här med att man pratar om det här
216 som ett nytt fenomen och någonting nytt, har jag svårt alltid svårt att ta till mig. Allting utvecklas hela tiden och hade
217 man kunnat ta den här informationen för 10 år sedan så hade man gjort någonting av det då också.

218

219 **M: Ja det är ju som du säger att det utvecklas hela tiden men att folk kanske inte förstår att de kan skapa**
220 **värde utifrån vad deras kunder gör på facebook eller vad de handlar på internet osv. Folk förstår inte att de**
221 **kan koppla deras kunders internetvanor till vart de resor och vad de handlar osv. Att de kan göra mer riktad**
222 **marknadsföring utifrån den här komplex datan som man kanske inte riktigt förstår. Men absolut, det är**
223 **som du säger att det bara är något som utvecklas kontinuerligt liksom**

References

- Arthur, L. (2013). *What Is Big Data?*. Available at: <http://www.forbes.com/sites/lisaarthur/2013/08/15/what-is-big-data/#1d80b2083487> [Accessed 15 May 2016].
- Barney, J. B. (1991). Firm resources and sustained competitive advantage. *J. Management* 17(1) 99–120.
- Bechmann, A. (2014). *Non-Informed Consent Cultures: Privacy Policies and APP Contracts on Facebook*. *Journal of Media Business Studies*
- Bekmamedova, N., Shanks, G. (2014). *Social Media Analytics and Business Value: A Theoretical Framework and Case Study*. 2014 47th Hawaii International Conference on System Science.
- Broadbent, M., P. Weill, B. S. Neo. (1999). *Strategic context and patterns of IT infrastructure capability*. *J. Strat. Inform. Sys.* 8(2) 157–187.
- Carlsson, S. (2016) *Åtta av tio styrelsemedlemmar saknar digital kompetens*. Available at: <http://digital.di.se/artikel/atta-av-tio-styrelsemedlemmar-saknar-digital-kompetens> [Accessed 15 May 2016].
- Chen, H., Chiang, R.H.L., Storey, V.C. (2012). *Business Intelligence and Analytics: From Big Data to Big Impact*. Special Issue: Business Intelligence Research, *MIS Quarterly* 36:4, 1165-1188.
- Cohen, W. M., D. A. Levinthal. (1990). *Absorptive capacity: A new perspective on learning and innovation*. *Admin. Sci. Quart.* 35(1) 128–152.
- Cole, D. Nelson, J. McDaniel, B. (2015) *Benefits and risks of big data*. SAIS 2015 Proceedings. Paper 26.
- Demchenko, Y., Grosso, P., Laat, C., & Membrey, P. (2013). *Addressing Big Data Issues in Scientific Data Infrastructure*. *IEEE*.
- Dosi, G., R. Nelson, S. G. Winter. (2000). *The Nature and Dynamic of Organizational Capabilities*. Oxford University Press, Oxford, UK.
- Duncan, N. B. (1995). *Capturing flexibility of information technology infrastructure: A study of resource characteristics and their measure*. *J. Management Inform. Sys.* 12(2) 37–57.
- Ekelund, A. (2016) *Svenska företag digitalt svaga*. Available at: <http://digital.di.se/artikel/svenska-foretag-digitalt-svaga-2> [Accessed 15 May 2016].
- Gantz, J., & Reinsel, D. (2011). *Extracting Value from Chaos*. IDC iView, 12.
- Gao, J., Selle, S., Koronios, A. (2015). *Towards A Process View on Critical Success Factors in Big Data Analytics Projects*. Twenty-first Americas Conference on Information Systems, Puerto Rico.
- Goes. P.B. (2014), *Big Data and IS Research*. *MIS Quarterly*, 38(3), iii-viii.

Groves, R.M. & R.L. Kahn (1979). *Surveys by Telephone: A National Comparison with Personal Interviews*, Academic Press, New York.

Halevi, G. and Moed, H. (2012). *The Evolution of Big Data as a Research and Scientific Topic: Overview of the Literature - Research Trends*.

<http://www.researchtrends.com/issue-30-september-2012/the-evolution-of-big-data-as-a-research-and-scientific-topic-overview-of-the-literature/> [Accessed 15 May 2016].

Hitt, M. Ireland, D. and Hoskisson, R. (2010). *Strategic management, Concepts, Competitiveness and Globalization*. South-western College Pub.

Hollison, M. (2015) *5 ways big data will change sales and marketing in 2015*. Available at: <http://www.inc.com/mick-hollison/5-ways-big-data-will-change-sales-and-marketing-in-2015.html> [Accessed 15 May 2016].

Jacobsen, D. I., G. Sandin and C. Hellström (2002). *Vad, hur och varför : om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen*, Lund : Studentlitteratur, 2002 (Lund : Studentlitteratur).

Keen, P. G. W. (1991). *Shaping the Future Business Design Through Information Technology*. Harvard Business School Press, Cambridge, MA.

Kiron, D., Ferguson, R.B., Prentice, P.K. (2013). *From Value to Vision: Reimagining the Possible with Data Analytics*. MIT Sloan Management Review: Spring Research Report (2013), 1-19.

Kohavi, R., Rothleder, N., Simoudis E. (2002). *Emerging Trends in Business Analytics*. Communications of the ACM 45:8, 45–48.

Kohli, R., Grover V., Business Value of IT (2008). *An Essay on Expanding Research Directions to Keep Up With the Times*. Journal of the Association for Information Systems 9:1 23-39.

Larson, K., Watson, R. (2011). *The Value of Social Media: Towards Measuring Social Media Strategies*, in Proc. of the 32nd International Conference on Information Systems, China.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. Available at: <http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation> [Accessed 15 May 2016].

Marr, B. (2014). *Big Data: The 5 Vs Everyone Must Know*. Available at: <https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know> [Accessed 15 May 2016].

Nelson, R., S. Winter. (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge, MA.

Neumann. P.G., (2014). *Risks and Myths of Cloud Computing and Cloud Storage*. *Communications Of The ACM*. October 2014; 57(10): 25-27. doi:10.1145/2661049.

Prahalad, C. K., G. Hamel. (1990). *The core competence of the corporation*. Harvard Bus. Rev. 63 79–91.

- Press, G. (2013) *A Very Short History Of Big Data*. Available at: <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/2/#b91df871af05> [Accessed 15 May 2016].
- Rider, F. (1944). *The scholar and the future of the research library*. New York city: Hadham Press.
- Rouse, M. (2013) *Big data management*. Available at: <http://searchdatamanagement.techtarget.com/definition/big-data-management> [Accessed 15 May 2016].
- Santhanam, R., & Hartono, E.. (2003). *Issues in Linking Information Technology Capability to Firm Performance*. MIS Quarterly, 27(1), 125–153.
- Saravana, C. (2013). *Big Data and Analytics – Value Chain Cross Section*. Available at: <http://www.business2community.com/big-data/big-data-and-analytics-value-chain-cross-section-0589031#Y3FFm5HvqZGHmkPr.97> [Accessed 15 May 2016].
- Sicular, S. (2012). No Data Scientist Is an Island in the Ocean of Big Data. Gartner.
- Sicular, S. (2013). *Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s*. Available at: <http://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/#1f9eaf343bf6> [Accessed 15 May 2016].
- Sorofman, J. (2013) *Data, Data Everywhere, Gartner Blog Network*. Available at: <http://blogs.gartner.com/jakesorofman/data-data-everywhere/> [Accessed 15 May 2016]
- Spears, J.L., & Barki, H. (2010). *User Participation in Information Systems Security Risk Management*. MIS Quarterly, 34(3), 503-A5.
- Wade, M., & Hulland, J.. (2004). *Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research*. MIS Quarterly, 28(1), 107–142.
- Watson, Hugh J. (2014). *Tutorial: Big Data Analytics: Concepts, Technologies, and Applications*. Communications of the Association for Information Systems: Vol. 34, Article 65.
- Weill, P. 1992. *The relationship between investment in information technology and firm performance: A study of the valve manufacturing sector*. Inform. Systems Res. 3(4) 307–333.
- Wernerfelt, B. (1984). *A resource based view of the firm*. Strat. Management J. 5 171–180.
- Westermeier, J.T. (2005). *Managing Open Source Software Risks in M&A Corporate Transactions*. Journal of Internet Law, 9(5), 20-24.
- Zhu, K., K. L. Kraemer. (2002). *E-Commerce metrics for net-enhanced organizations: Assessing the value of e-commerce to firm performance in the manufacturing sector*. Inform. Systems Res.
- Aureus Analytics (2015). *Infographic: Big Data And Analytics Trends For 2015*. Available at: <http://blog.aureusanalytics.com/infographic-big-data-analytics-trends-2015/> [Accessed 15 May 2016].

How Facebook advertising works. Available at:

https://www.facebook.com/business/products/ads/?campaign_id=1604904836426745 [Accessed 15 May 2016].

IBM (2013). *The four V's of big data*. Available at:

http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg [Accessed 15 May 2016].

International Accounting Standards Board (IASB)(2006). *Conceptual Framework Asset Definition*.

Available at: <http://www.ifrs.org/Current-Projects/IASB-Projects/Conceptual-Framework/Other-Public-Meetings-Observer-Notes/Documents/WSSAGENDAPAPER1A.pdf> [Accessed 15 May 2016].

SINTEF (2013). *Big Data, for better or worse: 90% of world's data generated over last two years*.

Available at:

<https://www.sciencedaily.com/releases/2013/05/130522085217.htm> [Accessed 15 May 2016].