Stockholm School of Economics
Department of Marketing & Strategy
Master Thesis, Fall 2017

In God We Trust - All Others Must Bring Data

A study on the adoption of machine learning among companies in Sweden

Linda Stoby Höglund (22366)

Abstract: The goal of this study is to investigate the organisational and managerial gaps impeding the successful adoption and deployment of machine learning in many organisations. The aim is to narrow the research gap on why the adoption of machine learning is slow, and provide further insights on what actions organisations must take to encourage advancements. The findings build on qualitative data from in-depth interviews with data scientists, analysts and analytics managers in 16 different organisations in Sweden. The results show that there are essential gaps in resources and capabilities which are delaying advancements of machine learning. This study make use of technology innovation theory and resource-based theory to analyse the identified gaps and suggest capabilities organisations should build to bridge them. The study concludes that two of the most important capabilities to develop to promote the adoption of machine learning is the ability to systematically educate management and business units, and the ability to connect analytics to strategy.

Supervisors: Per Andersson & Lei Huang
Examinator: Patric Andersson

Key Words: Big Data, Big Data Analytics, Artificial Intelligence, Machine Learning, Technology Innovation, Resource-based View

Acknowledgements

A warm thank you to Professor Per Andersson and Lei Huang for your guidance, advice and recommendations on how to progress with this study,

Thank you to all interviewees who have generously taken their time to share their experiences, thoughts and insights on this topic, and without whom this study would not have been possible,

Thank you to Katarina and Henrik for reading this thesis and suggesting improvements,

And finally, thank you to my family for your love and support.

"AI is the planet we're headed to. Machine learning is the rocket that's going to get us there. And big data is the fuel".

– Pedro Domingo, The Master Algorithm

Table of Content

Chapter 1. Setting the Scene ... 5
1.1 Introduction ... 5
1.2 Machine learning in a business context ... 6
1.3 Research motivation ... 7

Chapter 2. Establishing the research gap ... 8
2.1 Pre-study ... 8
2.2 Research gap and purpose ... 9

Chapter 3. Background ... 11
3.1 Big Data, Big Data Analytics & Machine Learning in Data Science ... 11
3.2 Introduction to Artificial Intelligence and Machine Learning ... 12
3.3 The challenges of adopting new technology ... 14

Chapter 4. Theoretical Framework ... 15
4.1 Technology Innovation Theory ... 15
4.2 Resources and capabilities ... 16
4.3 Resource-based theory ... 17
4.4 Technology Innovation Framework ... 19
4.5 Resources ... 20
4.6 Capabilities ... 22

Chapter 5. Methodology ... 24
5.1 Methodological fit ... 24
5.2 Data collection ... 24
5.3 Quality of Study ... 26

Chapter 6. Empirics ... 27
6.1 Identified Gaps ... 27
6.2 Resources ... 36
6.3 Capabilities ... 38

Chapter 7. Analysis ... 40
7.1 Gaps in Vision, Knowledge Development and Integration/application ... 40
7.2 Essential Resources ... 42
7.3 Essential Capabilities - Bridging the Gaps ... 42
7.4 Additional findings: The Importance of Middle Managers ... 46
7.5 Revisiting the analytical framework ... 47

Chapter 8. Conclusion ... 48
8.1 Theoretical Contributions ... 48
8.2 Managerial Implications ... 48
8.3 Limitations ... 49
8.4 Future Research ... 49

List of Works Cited ... 50

Appendix ... 56
Appendix 1: Big Data – The other four “Vs” ... 56
Appendix 2: Interviewee Guide ... 57
Appendix 3. Identified Gaps in Participating Organisations ... 58
Appendix 4. Interview Guide ... 59
Appendix 5. Advanced analytics ... 60
Appendix 6: Illustration of Data Management ... 61
Appendix 7: The Machine Learning Process ... 62
Appendix 8: Different Machine Learning Techniques ... 63

List of figures:

Fig. 1 The fields of data science

Fig. 2 The link between AI and big data

Fig. 3 A timeline of artificial intelligence

Fig. 4 The areas of machine learning and deep learning

Fig. 5 Example of application areas of machine learning in business

Fig. 6 Illustration of research process

Fig. 7 Different types of knowledge

Fig. 8 Innovation theory framework

Fig. 9 Technology innovation framework

Fig. 10 Analytical framework of resources and capabilities

Fig. 11 Summary of gaps

Fig. 12 Technology vs strategy driven approach for adopting machine learning

Fig. 13 Summary of capabilities to bridge identified gaps

Fig. 14 Comparison of framework

Fig. 15 Interviewee guide

Fig. 16 Top identified gaps in each industry

Fig. 17 Different types of analytical methods

Fig. 18 The analytics ladder

Fig. 19 The data management process

Fig. 20 The machine learning process

Fig 21 Different machine learning techniques

Fig. 22 AI index

Glossary

Advanced analytics: Although there is no universal definition of the term "advanced analytics" it generally refers to predictive and prescriptive analytics (appendix 5). These are the forms of analytics that will generate insights beyond standard business intelligence. The IT consultancy firm Gartner provides a comprehensive explanation, describing advanced analytics as "...examination of data or content using sophisticated techniques and tools…to discover deeper insights, make predictions, or generate recommendations. Advanced analytic techniques include those such as data/text mining, machine learning, pattern matching, forecasting, visualization, semantic analysis, sentiment analysis, network and cluster analysis, multivariate statistics, graphs analysis, simulation, complex event processing and neural networks" (Gartner, 2017).

Artificial Intelligence (AI): Is the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages (Oxford Dictionary, 2017).

Analyst: In the context of this study *analyst* refers to a data analyst. Data analysts can have a wide range of tasks. Traditional analysts work with business intelligence and descriptive analytics, but analysts can also refer to people working with big data and big data analytics. Analysts can either have a background in data science or a more traditional background in statistics, business or economics. The term analyst lends itself to some confusion, especially in the context of this study, as it can refer to people with no knowledge in big data analytics and machine learning, or people who are very well acquainted with the subject. In this study business intelligence analysts are referred to as 'traditional analysts' and analysts refer to people who have knowledge in big data analytics and machine learning.

Big data: Big data are data sets characterized by their volume, velocity and variety. Volume refers to the large quantity of data points in each set. Velocity refers to data being generated as a stream rather than in batches, and often in real-time. Variety refers to the different formats, text, images, audio etc. Recent research assign additional attributes to describe big data (see appendix 1).

Big data analytics (BDA): Because of the size of data sets, traditional analytics methods and software are unable to process big data to extract insights. With the help of today's technology new analytical tools are developing to manage big data. Big data analytics are methods to discover patterns, detect correlations and extract insights from big data sets. Machine learning is an example of a technique used in big data analytics.

Business Intelligence (BI): A collective term for the collection and analysis of data on business operations. It is common that BI reports on the firm's historical performance.

Chief Information Officer (CIO): The most senior executive in a corporation who is responsible for the firm's IT- technology, infrastructure and computer systems.

Data Management: Data management is a broad term that usually refers to the acquisition, recording, cleaning, integration and annotation of data. Data management is the process that proceeds data analysis (see appendix 6 for a further explanation).

Machine Learning (ML): A method to detect patterns in large sets of data, and applied in big data analytics. In machine learning computers learn from experience (input data) and generates output without explicitly being programmed on how to do it. It is a data analysis method where computers learn from data rather than hard coded rules

Visualisation: Refers to the visualisation of data in order to facilitate communication and understanding of what information the data holds. Usually involves statistical graphs, plots or information graphics.

Chapter 1. Setting the Scene

1.1 Introduction

In 2016, Donald Trump took home a startling victory in the US presidential elections. The unexpected win has partially been credited to Trump's successful online campaign. The company behind it, Cambridge Analytica, is a UK based big data company specialising in data driven political campaigns (Cambridge Analytica, 2017). In the aftermath of last year's election, the company gained worldwide attention for their "psychographic profiling" methods, through which they used extensive data on voters to create personality profiles. The data used on voters ranged from shopping data, bonus cards, membership cards, automotive data, social media and more. The different personalities created were then targeted with personalised content online, designed to affect them emotionally. The company has claimed to have had "...up to 5000 data points on over 230 million American voters" during the 2016 presidential race (Gregshorn, 2017). It is an exceptional large set of data. The analytical technique that made it possible to detect patterns and extract insights from such a huge amount of structured and unstructured data is machine learning.

Machine learning is also to blame for your Netflix addiction. In 2017, the streaming service had 110 million subscribers across 190 countries, and boosting over 14000 titles in its catalogue (Netflix, 2017). The company continuously collect data on what each subscriber watches, when they watch it, the device they are using, the recommended content the customer disliked or ignored, and what titles are the most popular. Algorithms powered by machine learning are continuously fed the data and the result is content designed to retain the viewer's attention. A similarity algorithm provide recommendations on movies akin to the one you just watched or liked. A personalized video ranker algorithm selects the order of the videos you are presented with in each genre row. Netflix knows it takes the average viewer less than 60 seconds to decide what to watch. Thus the order in which viewers are presented with videos is important. It has been reported that the recommendations feature on Netflix has cut customer churn with several percentage points and resulted in the company saving over $1 billion USD every year (Gomez-Uribe & Hunt, 2016). Netflix is an example of how big data analytics and machine learning drive business value.

Trump's election campaign and Netflix are only two examples of how big data analytics and machine learning algorithms are affecting our daily lives, and their presence is growing rapidly. The advancement of artificial intelligence and machine learning was named the most impactful tech trend in 2017, and is predicted to remain a top trend in the next couple of years (Rao, Voyles & Ramchandani, 2017). McKinsey Global Institute estimates that in 2016, corporations globally invested $26 to $39 billion in artificial intelligence, and the majority of it in machine learning (Bughin et al., 2017). The power of machine learning comes from the opportunities the technology offers to detect patterns and trends in vast amounts of data. It has the potential to unlock the value of information stored in big data, which traditional statistical tools have been unable to do. It opens up for more advanced analytics of unstructured data. Industry experts estimate that businesses which fail to take advantage of big data analytics will struggle to stay competitive in just three to five years (Lane, 2017). Yet, recent industry reports suggest that many organisations are struggling to adopt machine learning (Bedi et al., 2017). Further research is thus needed to investigate why the adoption of machine learning is slow, as well as what potential gaps organisations must bridge to effectively implement big data analytics and adopt machine learning.

1.2 Machine learning in a business context

Machine learning (a sub-area in AI) is a method to detect patterns in large sets of data - a capability with widespread application in business. Research has shown that organisations which are data driven outperform organisations that are not on objective financial and operational results (McAfee & Brynjolfsson, 2012). Thus, taking advantage of big data and the potential that machine learning offers to generate insights have competitive significance for businesses (Chen et al., 2012).

Finance & Banking	Retail & E-Commerce	Marketing & Sales	Travel & Boking	Health Care
Credit Scoring	Demand Forecasting	Market Segmentation	Demand Forecasting	Increase in diagnostic accuracy
Fraud Detection	Price Optimization	Churn Rate Analysis	Price Optimisation	
Risk Analysis	Recommendations	Customer life-time value prediction	Price Forecasting	Identify at risk customers
Client Analysis	Fraud Detection		Recommendations	
Trading & Forecasting	Customer Segmentation	Forecasting		

Fig. 1 Examples of application areas for machine learning in business

Because the application of AI and machine learning is still in its outset, there are a limited number of studies on how organisations relate and react to the developments in these areas. In the literature review of this study, three relevant reports were found. A study by McKinsey Global Institute (2017) surveyed 3073 chief executives globally about the deployment of AI in their organisations. The other study, conducted by Oxford Economics on behalf of Servicenow (Bedi et al., 2017) investigates the views of machine learning in corporations. The results are based on a survey among 500 Chief Information Officers (CIOs) across the globe. Both studies provide new insights on the adoption and deployment of AI and machine learning in organisations from a managerial perspective. A third study by the International Institute for Analytics share insights on the views of US companies on advanced analytics in comparison to traditional business intelligence (IIA, 2016).

The early AI adopters are large firms

The early adopters of AI are large firms (more than 500 employees) with digital awareness. The advantage that large firms have over small ones in adopting AI is access to large amount of data as well as human resources with the knowledge in data science and perhaps even machine learning. Organisations that have kept up with the digital evolution are generally better equipped and thus more prone to venture into AI development. According to the MGI digitalisation Index (appendix 9), high tech, telecom and automotive industries are most prone to adopt AI. Businesses in these markets already have a high level of IT resources and capabilities to build on. At the bottom of the digitalization index are health care, construction and travel/tourism industries. (Bughin et al., 2017)

There are differences in machine learning maturity

Although some industries might be more ready to adopt AI than others, there are differences at chief executive level as well. The ServiceNow (2017) study reveals that 53% of the CIOs recognise machine learning as a core priority, as their responsibilities grow from traditional IT operations to involve business strategy. The same survey also reveals that 69% of the CIOs thought decisions made by machines would be more accurate than those made by humans. More than half thought decision making facilitated by machines would be very important to their respective organisation within the next three years. The IIA report nevertheless concludes that although there is a strong interest in machine learning, there is a significant gap between perceived importance and performance. Firms state that exploring data to identify route causes, and building predictive models using machine learning techniques are considered important, but few are acting on it (IIA, 2016).

Europe is lagging and Sweden is falling further behind
In addition to variations in industries, there is also a significant differences in machine learning maturity between countries. The Americans display a stronger confidence in machine learning and expect greater value from the technology (Bedin et al., 2017; Bughin et al., 2017). Europe lags behind both North America, Asia and the Pacific in driving digital transformation through machine learning. The European CIOs believe the technology is important but they are not making the same level of investments. They exhibit lower expectations on how machine learning could improve their businesses. In Europe Sweden, along with the Netherlands, are slowest in advancing machine learning. Of the 46 surveyed Swedish CIOs only 28% indicate that machine learning is a strategic focus for them, significantly lower than for the other European countries participating in the study. An interesting observation is that the Swedish CIOs rank the challenges to machine learning adoption - 'insufficient data quality', 'outdated processes', and 'lack of budget for new skills' - significantly lower than their European peers. Unfortunately the study offers no further insights on why Sweden appears to be lagging behind in advancing machine learning despite perceiving the challenges as significantly lower than the European countries leading the development.

1.3 Research motivation

Previous research establishes that analytics provide firms with a competitive advantage. Machine learning methods make it possible for organisations to embed advanced analytics in their operations to improve performance. If the benefits of advanced analytics are clear, why are not all firms and organisations deploying machine learning? What factors are holding organisations back? The reports by McKinsey Global Institute and Service now provide some insight into the organisational and managerial challenges. There is nevertheless a need to further understand the internal gaps organisations must address to adopt machine learning. Advanced analytics and machine learning hold great value potential for firms, but in order to progress they must eliminate any stumble blocks for innovation. The goal of this study is thus to contribute with valuable insights on what gaps organisations must bridge, and what capabilities they must develop to do so.

Chapter 2. Establishing the research gap

Existing literature and reports suggest that more research is needed to understand the underlying reasons to why adoption of machine learning is slow in some organisations. To get a better grasp of the current situation an explorative pre-study was conducted.

2.1 Pre-study

The purpose of the explorative pre-study was to get a 'sense' of the problematics surrounding the adoption of machine learning. Three in-depth interviews were conducted with a sales director in a large telecom cooperation, a data scientists/IT-consultant specialising in machine learning, and a senior solutions consultant at large technology cooperation. All interviewees have experience in implementing machine learning projects and were able to share insights on the process, as well as give their perspective on significant organisational and managerial gaps. The key findings from the pre-study were:

- The limitations to implementing machine learnings are not related to technology itself. The technological tools are available, but the limiting factors relate to the organisation (company culture, office politics, organisational structure, insufficient resources etc.) and human factors (lack of knowledge, interests, influence, support etc.).

- There is limited use of data in general in many organisations, or a scepticism towards data analytics. There is little or no advocacy for the advancement of big data analytics.

- Adopting machine learning involves organisational changes, and there is an inherent inertia in making changes to set organisational practices, routines and structures.

"*We know that we need to adopt new analytical tools to become more data driven, but it is a huge organisational challenge. Data exists in so many different forms and databases across the organisation. We are already divided and to create a unified system we need to establish new paths of communication between different departments which are non-existent at the moment…"*
– Sales Director, Telecom

- There is a lack of knowledge of what the technology can bring and how to implement it. It is a new phenomenon and although it has been hyped in the media, many managers don't know what to do with that information.

"It is a problem that management often lack a fundamental understanding for the technology - the limitations and the possibilities. As a data scientist in machine learning, management expects you to solve all their problems and make them data-driven, but it is not a simple process. They don't understand what tools and resources you need from them…"
- IT Consultant

2.2 Research gap and purpose

Big data has been declared a major economic resource and leading tech companies are already reorienting themselves around artificial intelligence (Dong, 2017). Big data analytics is expected to generate competitive advantages that will make or break corporations in the next two to five years (Lane, 2017). Machine learning is at the forefront of these digital developments.

However, recent observations suggest that many organisations struggle with the implementation of machine learning. Supported by the observations of existing reports, the pre-study illustrate gaps in organisational and managerial resources which hampers the adoption of machine learning. Further research is required to investigate what gaps are present and what organisations can do to bridge them in order to effectively adopt and deploy machine learning.

The **purpose** of this study is to investigate what the organisational and managerial gaps are and how organisations can bridge them to successfully adopt and deploy machine learning. The aim is to narrow the research gap on why adoption is slow in many organisations, and provide further insights on what actions organisations must take to encourage advancements.

This study strives to answer the following **research questions**:

1. *What are the organisational and managerial gaps delaying the adoption of machine learning in organisations?*

2. *What actions must organisations take to enable the adoption of machine learning?*

Research outline

To address the identified research gap, a qualitative study was conducted. To answer the research questions, a deeper insight into the phenomenon was needed. Because the application of machine learning in a business context is rather new, especially in Sweden, it was not possible to find a single case study that could provide enough research depth into the subject. Thus, a multiple-case study design was used to study the commonality of identified gaps, and achieve research depth by comparing and contrasting cases (Bryman & Bell, 2007). An abductive approach, known as "systematic combining" was chosen, as it allowed the author to freely move between theory and empirics to determine research direction and scope (Dubois & Gadde, 2013).

The study was divided into two section: a pre-study and a main study. The purpose of the pre-study, as described in qualitative research methods was to (1) expand the author's understanding of the subjects and potential conflicts, (2) understand the wider context of the subject, and (3) confirm the research gap, and study scope (Flick, 2014). Based on the pre-study, the research questions were defined and the main study designed accordingly. The study combines primary and secondary data sources. The primary sources include interviews with professionals in various industries (for a full list see appendix 2). The interviews were also supplemented with expert interviews (consultants in the field). The author further attended a full-day industry conference with lectures and discussions on the topic of data analytics to gather first-hand information on the latest news in the subject field. Secondary data was collected through academic journals, news articles, industry journals and company websites.

Delimitations

Artificial intelligence and machine learning are rather 'hyped' subjects at the moment. It is often said that AI is something "everybody is talking about but nobody is doing" (Analyticsdagarna, 2017). It is an area that lends itself to different research approaches. Due to the author's background in business and management, the study was limited to understanding the managerial and organisational gaps, rather than dwelling deeper into the technology (this also seemed appropriate as the pre-study indicated that the major challenges in adopting machine learning are limitations in management and organisation, rather than technology). The scope of case study companies was limited to Sweden because (1) previous reports indicate that companies in Sweden are lagging behind in the adoption of machine learning, and (2) it allowed the author better access to conduct interviews. The companies approached for a case study were prioritised depending on (1) how far along they had come in adopting machine learning, (2) size, (3) business maturity and (4) industry. Chosen companies had at least begun the process of adopting machine learning, which made them able to share insights on the process. The companies also had a reasonable sized organisation and had been in business for some time (no start-ups). The author made a judgment on a case-to-case basis to determine if the company sufficiently fulfilled the criteria. Finally, consideration was given to industry and finding appropriate cases in industries with different level of AI maturity. The AI index (Bughin et al., 2017) was used as a guide to determine in what industries to look for cases. The decision was made to pick some industries from the top, some from the middle and some from the bottom of the AI index (see appendix 9 for index illustration).

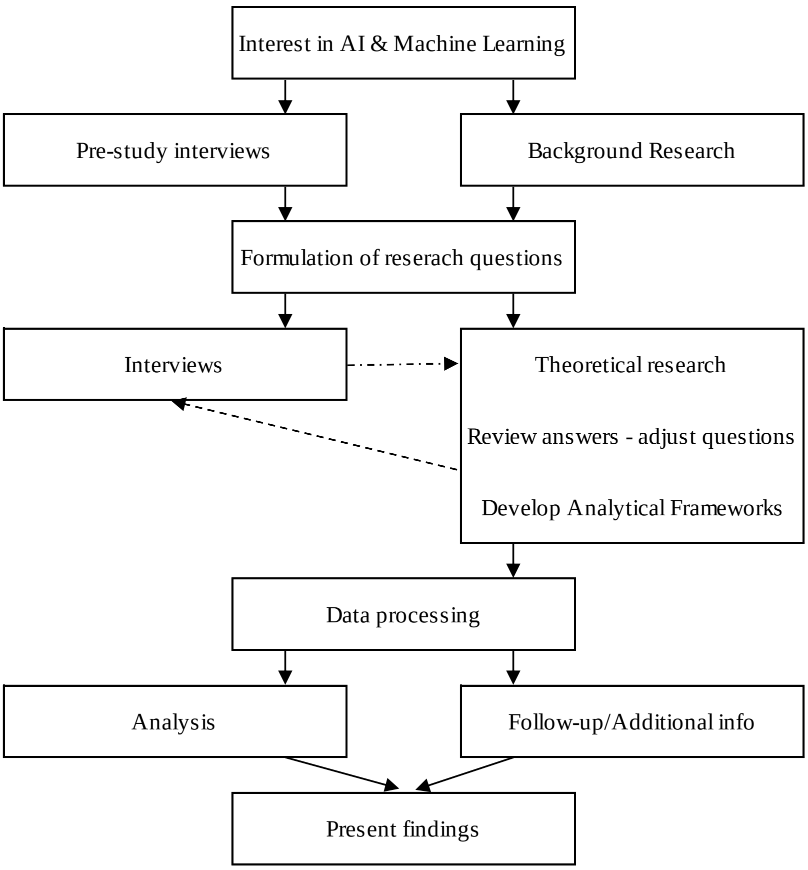

Fig. 2 Illustration of the research process. Dotted arrows suggest iteration of this process, as part of the abductive method where the researcher moves between theory and empiric results from interviews

Chapter 3. Background

In order put the purpose and research questions of this study in context, this chapter offers a brief explanation of artificial intelligence, machine learning, big data and how these subjects relate to one another. The chapter ends with a retrospective view on technology adoption and the organisational challenges of computerisation that many corporations experienced in the late 90s.

3.1 Big Data, Big Data Analytics & Machine Learning in Data Science

Big data, big data analytics and machine learning are all sub-fields of data science. Data science, or data-driven science, is defined as "...an interdisciplinary research field about scientific methods, processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured" (Dhar, 2013).

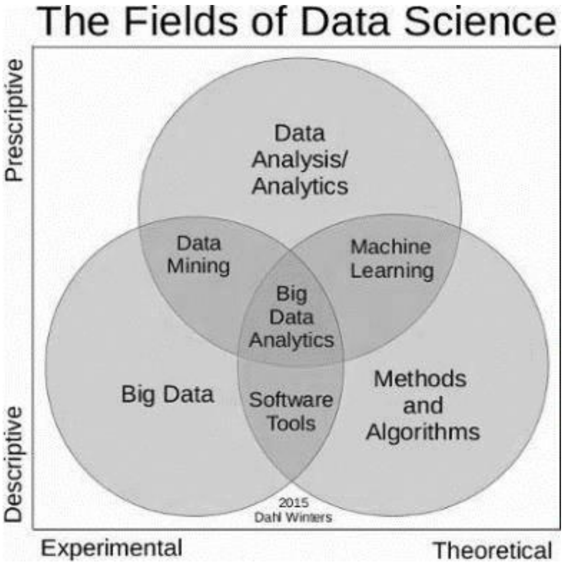

Fig. 3 The fields of data science (Winters, 2015)

The Evolution of Big Data Analytics

The importance and potential of big data and analytics in supporting strategic goals are widespread notions in business (Davenport 2006; Manyika, et al., 2011; McAfee & Brynjolfsson 2012). The advancements in digital technology and connectivity allow organisations to collect vast amount of data in virtually every business area and function. In the past, data analytics have relied on Bayesian statistics and econometric models to extract information and insights to aid businesses in their decision making. However, as the *volume* of data has increased from terabytes to petabytes, *velocity* has gone from snapshots to high-frequency streaming data, and the *variety* has expanded to include, numeric, network, text, images and video, traditional analytical models are now insufficient (Galbraith, 2014; Wedel & Kannan, 2016). Consequently there has been a growing demand for new analytical methods that can handle big data.

The term big data analytics (BDA) was originally coined in 2012 (Chen, Chiang, & Storey) and has since been defined and described in numerous ways by different authors. Lamba and Dubey (2015) define big data analytics as "...the application of multiple analytical methods that address the diversity of big data to provide actionable descriptive, predictive and prescriptive results"(p.5).

Emerging literature argues that big data analytics has a significant positive effect on firm performance, and that it is becoming an integral aspect of businesses' decision-making processes (Germann, Lilien, Fiedler & Kraus, 2014). It has furthermore been described as the "next frontier for innovation, competition, and productivity" (Manyika et al., 2011, p. 1). In retail, firms can leverage big data to improve customer experiences, personalise offerings and make just-in-time recommendation (Tweney, 2013). In manufacturing and supply chain management, big data analytics can enable optimisation of processes, cut costs and provide better monitoring (Davenport et al., 2012). In the health care sector, big data analytics can aid in predicting the outcome of treatments and improve patients' quality of life (Liu, 2014). There are many areas of application, and the businesses that master big data analytics can expect significant advantages.

3.2 Introduction to Artificial Intelligence and Machine Learning

Artificial Intelligence – the science of intelligent machines

The term artificial intelligence (and subsequently the research field) was coined by researchers in 1955 and described as "...to find how to make machines use language, form abstractions and concepts, solve the kinds of problems now reserved for humans, and improve themselves…"(McCarthy et al., 1955). Although there is no universal definition of artificial intelligence, more recent ones describe it as "...the science and engineering of making intelligent machines, especially computer programs…" (Hatch, 2012, p. 9) and "the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages" (Oxford Dictionary, 2017). As the latter definition suggests, artificial intelligence have many branches and sub categories. From this point on, a reference to AI in this study will indicate weak (also known as narrow) AI technologies, as these systems are the ones advancing today and are relevant in a business context. Weak AI systems carry out tasks based on set rules with a given type of information. An example is IBM Watson which, despite being inspired by human reasoning, generates its intelligence by mining huge quantities of information (data) using machine learning.

Linking Big Data and Artificial Intelligence

The extensive interest in artificial intelligence and machine learning is linked to the potential it holds in big data analytics. Innovation in data management and analytics have progressed, thanks to developments in AI and machine learning.

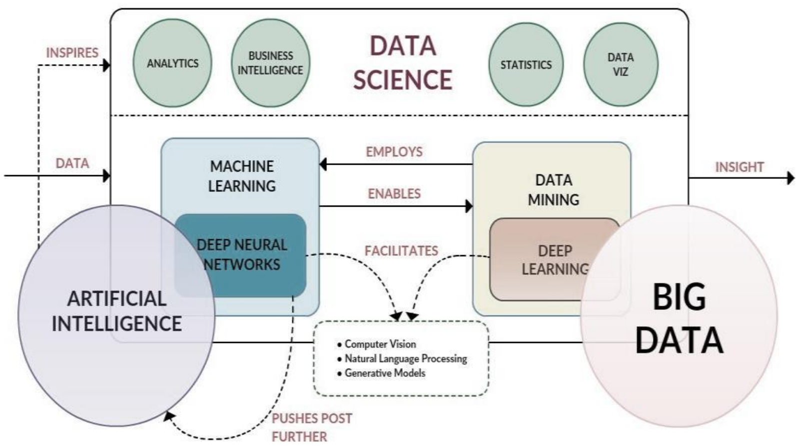

Fig. 4 The link between AI and big data ('What's the big data', 2016)

Corporations are investing in artificial intelligence

Artificial intelligence has been a buzzword in business for several years now, and significant investments are being made in AI technology. The early adopters of AI technology are found in tech, telecom, automotive and the financial services industries, while examples of lagging industries are education and health care. It is nevertheless the tech giants and digital native companies like Amazon, Apple and Google that are leading developments, putting an estimated $18 to $27 billion in internal corporate investments towards artificial intelligence (Bughin et al., 2017). In Sweden, companies such as Stena Line, and Scania have publicly announced their decision to explore the possibilities of AI (Rosengren, 2017) This interest is echoed by the Swedish government agency Vinnova, responsible for administering state funding for research and development. In 2017 the agency is funding more projects than ever before relating to artificial intelligence, **machine** learning and the Internet of Things (Analyticsdagarna, 2017).

Machine learning – a technique to extract information from big data

Machine learning, a form of AI, refers to "...the automated detection of meaningful patterns in data" (Shalev-Shwartz & Ben-David, 2014, p. vii). The term learning can be described as the process of converting experience into knowledge and expertise. In machine learning computers learn from experience (input data) and generate output without explicitly being programmed on how to do it. It is a data analysis method where computers learn from data rather than hard coded rules. At the core of machine learning are models of algorithms which iteratively learn from data to improve in accuracy. The iterative learning process is key to machine learning as it allows the models to independently adapt to new data, and makes it possible for computers to find hidden insights without being programmed for where to look. Instead they learn from previous computation to generate repeatable decisions and results (SAS, 2017). While machine learning algorithms have existed for decades, the ability to automatically apply complex mathematical calculations to big data is a recent development (SAS, 2017). Machine learning is a vast and complex area of study and although fascinating, the details of the technology is beyond the scope of this study (see appendix 7 for illustrations of machine learning processes).

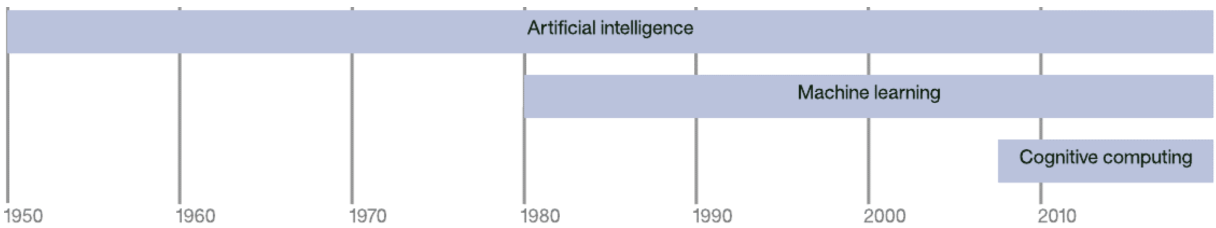

Fig. 5. A timeline of artificial intelligence (Jones, 2017)

Recent advancements in machine learning is related to neural networks

The most recent developments in machine learning are related to *deep learning* and *neural networks*. Deep learning architectures are applied in computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design etc. The potential of deep learning architectures are reflected by the investments many companies have made to advance the technology. In 2014 Google acquired the British start-up DeepMind for $650 million, which had created a neural network that could play videogames similarly to humans (Gibbs, 2014). In August 2017 Apple announced that it is going to use deep neural networks for its personal assistant Siri in iOS10 and upcoming iOS11, as it will allow smoother and more natural conversations with her (Siri Team, 2017). Deep learning is also the key to self-driving vehicles and in addition to Apple, car manufactures such as Toyota and Tesla have made significant investments in developing the technology (Bughin et al. 2017).

It can be concluded that the rapid development of various AI applications is derived from a breakthrough in deep learning and neural networks. Again, the timing of this development is not coincidental. The neural networks are fed terabytes of data to train the models. It could involve a year's worth of speech samples or hundreds of thousands of images, to make the computer learn how to recognise objects, words or sentences. This amount of data is now readily available and computers have the power to processes it, and thus deep learning is gaining momentum. An illustrative example of the exponential acceleration of deep learning is the rapid increase of deep learning projects pursued by digital natives such as Google. In 2012, the company announced that it was pursuing two deep-learning projects, four years later the number of deep learning projects at Google had increased to over a thousand (Parloff, 2016). In 2011 IBM's AI Watson defeated three human players in a remarkable game of Jeopardy. At the time IBM Watson was not using deep learning. Only five years later, deep learning has been used to augment almost all of Watson's inherent services (Jones, 2017).

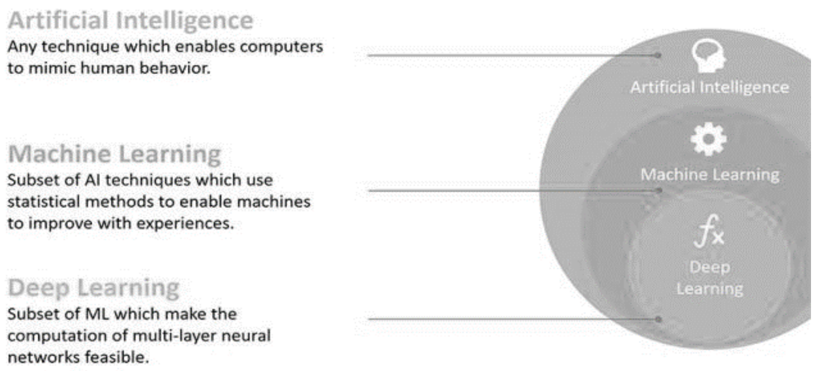

Fig. 6 The areas of AI, machine learning and deep Learning (Mierswa, 2017)

3.3 The challenges of adopting new technology

The challenge of effectively incorporating new technological advancements and capitalise on technological development, is not new. At the start of the new millennium computers were becoming more powerful and cheaper. At that time, researchers nonetheless concluded that the business value of computers was limited less by their computational capability and more by the ability of business managers and organisations to adapt processes and procedures to leverage this capability (Brynjolfsson & Hitt, 2000). Since first being introduced in organisations, investments in information technology has resulted in complementary changes in other parts of the organisation. There has always been an inherent challenge in matching organisational structure with technology capabilities (Brynjolfsson, Renshaw and Van Alstyne, 1996). Previous case studies on the adoption of new technology indicate that embedded routines, thinking-patterns and limited knowledge often hinder the transition to new technology (Brynjolfsson, Renshaw and Van Alstyne, 1996; Murnane, Levy and Autor, 1999). These observation in previous research support the results of the pre-study that the foremost challenges of adopting machine learning lays foremost with management and the organisational structure.

Chapter 4. Theoretical Framework

In this chapter the author draws on innovation theory and resource-based theory to understand how the identified gaps in knowledge, resources and capabilities affect organisations' ability to innovate and their potential to adopt machine learning. By reviewing existing literature and previous research two analytical frameworks are constructed – one for each research question.

4.1 Technology Innovation Theory

In the pre-study, the interviewees argue that a significant hurdle to the adoption of machine learning is the gap in knowledge. Many people across the organisation lack an understanding for big data, advanced analytics and machine learning. Existing literature on innovation emphasise the importance of knowledge in innovation – describing innovation as "…the application of knowledge to produce new knowledge" (Drucker, 1993, p. 173). From this theoretical perspective, knowledge is considered a key resource in driving innovation. Building on knowledge-management, Johannessen et al. (1999) further develop the notion of knowledge as instrumental in innovation by designing a theoretical framework on organisational innovation. The authors argue that the three factors affecting organisational innovation are:

(1) Vision: Vision is the tension between actual performance and desired future performance. An organisation's vision is largely determined by its existing market activities and knowledge. New technologies in markets where the organisation is not active, are thus often neglected due to "bounded vision." Bounded vision limits the strategic options in an organisation's view field, often resulting in underinvestment of new technology that sits in the periphery (Fransman, 1990).

(2) Knowledge creation: Innovation is also affected by an organisation's ability to create new knowledge. There are four types of knowledge: *explicit knowledge*, *systematic knowledge*; *relationship knowledge*; and *tacit knowledge*. There is a difference in how easy it is to communicate and apprehend a specific sort of knowledge (see fig. 7). By combining human resources – individuals in the firm - with different knowledge sets, new knowledge can be created (Johannessen et al., 1999).This exchange of knowledge and interactive learning is considered essential, and innovative organisations have proven to have highly effective learning systems (Tusman & Nadler, 1986). Individuals who have both theoretical and practical knowledge are able to see how their knowledge branch fits into the larger picture, and can collaborate with other knowledge branches to expand overall capabilities (Leonard-Barton, 1995).

	Attainable and easy to comprehend	Attainable but difficult to comprehend
Relatively easy to communicate	Explicit knowledge	Systemic-knowledge
Difficult to communicate	Relationship-knowledge	Tacit knowledge and hidden knowledge

Fig. 7. Different types of knowledge (Johannessen et al., 1999)

(3) Knowledge integration and application: for knowledge to contribute to innovation, it has to be spread throughout the organisation. Explicit and systematic knowledge is easy to communicate and share with others. Implicit and tacit knowledge, on the other hand, is harder to vocalise and transfer. Tacit knowledge is learned by using, doing and experimenting. Unlike explicit knowledge, which is often shared in a systematic way (i.e. through formal education), tacit knowledge is spread informally through socialisation (Stewart, 1997). To innovate, organisation cannot only rely on explicit knowledge. They must develop processes to ensure tacit knowledge is communicated and integrated into operations as well.

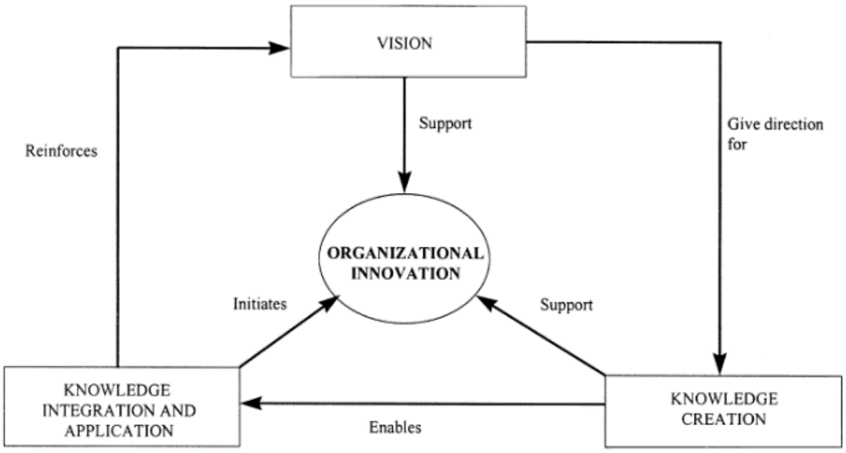

Fig. 8. Innovation theory framework (Johannessen et al., 1999)

Against this theoretical background it is helpful refine the first research question to:

***What are the gaps in vision, knowledge creation and knowledge integration/application delaying the adoption of machine learning*?**

4.2 Resources and capabilities

In addition to a lack of knowledge, the pre-study revealed that organisations lack other essential resources to effectively adopt machine learning, such as appropriate IT-infrastructure and quality data. Existing theory holds that organisations achieve sustainable competitive advantage by developing and deploying internal resources and capabilities (Grant, 1991; Peteraf, 1993; Wernerfelt, 1984). Resources are productive assets the firm own, while capabilities refer to what the firm can do (Größler and Grübner, 2006; Ulaga & Reinartz, 2011). In organisational structures, resource are transformed into capabilities which generates a competitive advantage. A firm's organisational capability can be defined as a "firm's capacity to deploy resources for a desired end result" (Helfat and Lieberman, 2002). Consequently, capabilities cannot be purchased, but must be built (Teece et al., 1997). One of the most dominant paradigms in strategy is the resource based-view which focuses on a firm's internal resources and capabilities to provide a competitive advantage.

4.3 Resource-based theory

The most frequently theoretical foundation applied in investigating the business value of IT and competitive advantage is the resource-based theory (Barua, Kriebel, & Mukhopadhyay, 1995; Bharadwaj, 2000; Mata, Fuerst & Barney, 1995; Melville, Kraemer, & Gurbaxani, 2004). Resource based theory (RBT) holds that a firm has a bundle of valuable resources that it applies to gain a competitive advantage (Grant 1991; Penrose, 1959). There are physical capital (tangible), human capital and organisational capital (intangible) resources (Barney 1991). Not all resources, however, are "strategic resources." In order to provide a strategic advantage resources have to be valuable, rare, imperfectly imitable and exploitable by the organisation (Barney, 1991; Lee & Grewal, 2004). A valuable resource improves the firm's bottom line results or offers customers something that competitors cannot. A rare resource is not abundant and imperfectly imitable, indicating it is difficult to copy. A particular company culture or exceptional leadership contributing to a firm's market position would be examples of resources that are scarce and difficult to imitate. An exploitable resource allows the firm to utilise it in a way that competitors are unable to (Barney, 1991).

In the context of big data analytics and machine learning, tangible resources would include computer servers, software and platforms used to collect, store and analyse data. Big data analytics and machine learning require a high level of computational power. Traditional software is unable to cope with the volume, velocity and variety of the data and requires a certain level physical capital resources (Bharadwaj, El Sawy, Pavlou, & Venkatraman, 2013; Davenport, Barth, & Bean, 2012). The human resources required to gain a competitive advantage from big data and machine learning are data scientists with the ability to extract insights and information from the data. They have the skill and knowhow on how to develop and apply the necessary tools to turn data into knowledge. Finally, organisational resources provide the indispensable framework in which the physical capital and human capital operate. It is the organisational structure that enables a firm to turn insights into action. Excellent computer power and talented data scientists can only provide the groundwork for big data analytics and machine learning. Realising the potential of advanced analytics and generating a sustainable competitive advantage comes down to internal business processes and management's ability to capitalise on the insights generated.

Resource based view - theoretical elements

There are three key elements to the resource-based view: resource functionality, resource creation and decay, and resource combination (Lockett, Thompson & Morgenstern, 2009).

(1) Resource functionality

Resource functionality emphasises that how a resource is put to use (its functionality) matter more than the resource per se (Penrose, 1959; Peteraf and Bergen 2003). Resources hold potential value and can be applied across a number of different functions, thus organisations must decide on the most profitable usage of the resources they control.

(2) Resource creation and decay

It is argued that a firm's resources are directly related to its past activities (Penrose, 1959). Firms engage in a range of activities which lead to the development of firm specific resources over time. From an RBV perspective, the growth of firms is a result of excess resources and capacity. The excess of resources provides a basis for expansion (Locket, Thompson & Morgenstern, 2009; Penrose 1959).

(3) Resource combination

Resource combination reference the fact that resources are more valuable in combination with other resources (Penrose, 1959). Competitive advantage rarely depends on a single resource, but comes down to a mix of resources. Resources can be complementary, related or co-specialised to one another. The notion of combining resources is central to understanding capabilities as a capability can be thought of "...the simultaneous deployment of resources and factors of production" (Teece et al., 1997). The capabilities of an organisation thus depends on how it opts to combine its resources. Capabilities arise from a specific bundles of resources. As demands change, organisations stay competitive by adding or redesigning bundles of resources to develop new capabilities (Lockett, Thompson & Morgenstern, 2009).

From this perspective, firms grow by identifying new market opportunities and subsequently change their combinations of resources to match the demand. Three different degrees of recombination activities have been conceptualised in previous literature: stabilizing, enriching and pioneering (Sirmon et al, 2007). The purpose of *stabilizing* activities is to maintain a certain level of competitive advantage by making minor incremental improvements to existing capabilities. It can, for instance, involve mandatory training of staff to ensure their knowledge and skills sets are up to date.

Enriching refers to the bundling of resources with the purpose to extend and elaborate current capabilities (Sirmon et al., 2007). Enriching actives go beyond keeping skills up to date, and strive to extend current capabilities by adding complementary resources to the bundle. The complementary resource could be part of the firm's resource portfolio already, or be newly acquired. Enriching activities can result in a competitive advantage, but in the long-term perspective capabilities that are only extended are likely to be imitated by competitors (Sirmon et al., 2007). To stay competitive over time, firms need to develop new capabilities, through *pioneering* activities (Sirmon et al., 2007). Pioneering actives are challenging and time consuming for firms, because instead of building on existing knowledge it requires exploratory learning. The purpose of pioneering is to build new competitive advantages by adding resources to the firm's portfolio and create new bundles to generate unique capabilities (Sirmon et al., 2007).

Pioneering demands a broad and deep knowledge base to detect and foresee the potential of developing advantageous capabilities by bundling resources that at first may seem unrelated (known as bisociation) (Di Gregorio, 2002). It requires a high level of creativity, and significant understanding for the functionality of a resources in order to appreciate its potential to create a novel capability when combined with unrelated resources. Uncertain or developing environments often require new capabilities, forcing firms operating in dynamic settings to continuously engage in pioneering bundling activities to stay competitive. Today's rapid digital evolution challenges organisations to acquire new resource and develop new capabilities to stay competitive. The RBV thus offers an interesting theoretical perspective to study how firms have adjusted (or not adjusted) their portfolio of resources and capabilities to adopt and deploy machine learning in data analytics.

Against this theoretical background it is helpful to refine the second research question to:

What resources and capabilities should organisations develop to bridge the identified gaps?

Constructing the Analytical Framework

Part I: The first part of the analytical framework outlines essential capabilities related to vision, knowledge creation and knowledge integration.

4.4 Technology Innovation Framework

Vision

Previous research shows that organisations have a tendency to inadvertently ignore new technologies outside their market activities (Johannessen et alt., 1999). However, because of the ‘hype’ surrounding artificial intelligence and machine learning, it is likely that most organisations have heard of the technology. The interesting aspect to determine is thus if the organisation has taken any action towards *recognising big data analytics and machine learning in their long-term strategies*, as an indication that machine learning is genuinely a part of the organisation’s future vision.

Knowledge Creation

Organisations that are highly innovative have effective learning systems, thus a *systematic learning system on how to exploit big data analytics and machine learning opportunities* would benefit adoption. Organisations apt at knowledge creation also enables the exchange of both explicit and tacit knowledge (Johannessen et al. 1999). Thus *creating opportunities for knowledge-exchange between data scientists and between data scientists and the rest of the organisation*, including management, is an important capability.

Knowledge Integration and Application

Tacit knowledge is considered valuable in innovation, but it is difficult to communicate such knowledge. Technical knowledge is often highly tacit (Teece, 1994). Machine learning, for instance, requires a lot experimenting. Consequently, each data scientist have unique experiences and knowledge that is valuable. To facilitate the adoption of machine learning it is thus that organisations have *strategies for how to integrate tacit knowledge in the organisational knowledge portfolio*.

Technology Innovation Framework
Vision
Recognising BDA and ML in long-term strategies
Knowledge Creation
Creating systemetic learning systems
Creating opportunites for exchanging knolwedge
Knowledge Integration & Application
Able to integrate tacit knolwedge

Fig. 9 Technology innovation framework

Part II: The second part of the analytical framework draws on existing research in adjacent fields to construct a framework of the essential resources and capabilities to adopt machine learning.

4.5 Resources

Physical resources

Data quality and quantity

Big data analytics and machine learning require large sets of data. Previous research indicate that IT-strategists and data analysts show a growing concern for the quality of data they are using (Brinkhues et al., 2014). Data quality is key to derive accurate insights and as datasets grow, quality in terms of completeness, accuracy, format, timeliness, reliability and perceived value becomes an important resource (Mikalef et al., 2017; Ren et al., 2016). Data availability is another essential resource. Existing literature indicate that combining data from different sources, even within the same organisation, can prove difficult because of existing IT architectures where information are kept in separate silos (Douglas, 2013; Fosso Wamba et al., 2015).

IT-infrastructure and hardware

Large unstructured datasets and the continuous inflow of data require IT architecture and process power to efficiently collect and store big data. It requires investment in new technology and increasing computer and server capacity to handle the great volume and variety of data (Gupta and George, 2016).

Information system and software

Existing literature in information systems indicate that new software that allows processing of unstructured data in a continuous flow, rather than in batches, is needed for big data analytics. The complexity of the data pushes the limits of current analytical software and tools (Douglas, 2013; Fosso Wamba et al., 2015).

Organisational resources

Coherent data governance system

Literature on big data analytics define data governance as “the approach that analytic-based organisations use to define, prioritize, and track analytics initiatives, as well as to manage different types and categories of data related to analytics…” (Espinosa and Armour, 2016). Adequate data governance is described as a key resource in research on big data, and a lack thereof is noted as a major hurdle in leveraging data (Garmiki et al., 2016; MIT Sloan, 2016; Posavec and Krajnovic, 2016). A fragmented structure makes it impossible capitalize on the potential different data sets have collectively. Data analytics is cross-functional and does not only require a coherent management system but also increased collaboration across departments and internal organisational boundaries. Data governance is an organisation-wide challenge and previous studies have found that effective data governance must be initiated and overseen by top management (Vidgen et al., 2017). It further requires management's commitment to data-driven decisions.

A data-driven culture

A data-driven culture can be defined as “an operating environment that seeks to leverage data whenever and wherever possible to enhance business efficiency and effectiveness” (Dykes, 2017). In a data driven culture, data is considered a core function of the organisation much like finance, sales or marketing (Grossman & Siegel, 2014). In this type of environment, machine learning becomes a natural mean to an end. Advanced analytics govern strategic decisions and machine learning is a technique to achieve the best possible outcomes to base decisions on. A data-driven culture has proven to influence the overall success of big data projects (LaValle et al, 2011).

What sets a data-driven culture apart is the integration of data in strategic decision making. Organisations can use data to improve virtually any aspect of the business, but few use analytics in a highly intentional manner to address or anticipate strategic challenges (LaValle et al., 2011; McAfee & Brynjolfsson, 2012; MIT Sloan, 2016). Existing literature shows that even though firms engage in big data analytics projects, a majority rely on managerial experience and intuition in strategic decision-making (Provost and Fawcett, 2013).

A fundamental challenge to overcome in creating a data-driven culture is the establishment of trust and acceptance for data-driven results across the organisation. Data driven decisions cannot be limited to top-management, middle-management or a single department. Making-decisions based on information and insight from data must become widespread practice (Gupta and George, 2017). Previous research shows that investing in big data analytics and a fact-based operating culture benefit the adoption of a data-driven culture (Lamba and Dubey, 2015; Kamioka and Tapainen, 2014). Case studies also demonstrate that data-driven organisations have leaders who are convinced of the importance of analytics in strategic decisions, and are confident in data generating competitive advantages (MIT Sloan, 2016).

Adequate organisation of competence

To effectively adopt and deploy machine learning, knowledge about the new technology has to be integrated with knowledge about the business problems it is expected to solve. It requires an organisation and bundling of resources that allow enriching and pioneering activities. Existing literature on organisational design demonstrates the benefits and drawbacks with both centralizing and decentralizing competencies in data science and analytics (Grossman and Siegel, 2014). Centralizing competencies make it easier to achieve a critical mass of talent and create a large pool of expertise on how to deploy advanced analytics and data mining models (Grossman and Siegel, 2014). It can be particularly valuable to have a “centre of expertise” when developing a new capability. Centralising competence nevertheless removes data scientists from the business units they are supposed to support.

A decentralized model allow analytics staff to be closer involved with business operations. It gives them a deeper understanding of core issues which allow them to combine their technical expertise with business insights to create new capabilities tailored to solve a specific business problems. Placing a smaller number of data scientists in each business unit resolves the lack of integration, but it also generates other challenges. Achieving critical mass on issues concerning the organisation at large, as well as dispersing and possibly diluting expertise are two problems generated by a decentralised model. Smaller units increase the pressure on data scientists’ skills. A smaller team run the risk of lacking knowhow (Berner, Graupner, & Mädche, 2014). Competence in analytics and machine learning must thus be organised in a way that it fosters a further expertise in the technology without losing insight into business operations.

Human resources

Top Management Engagement

Top management has the power to instigate initiatives and influence decisions that determine the future direction of an organisation. The successful adoption and deployment of machine learning can involve altering the organisational culture to become more data driven, hire a critical-mass of data scientists, invest in new IT infrastructure, or allow analytics to drive strategic decisions - all of which require top management's involvement. Existing literature in big data emphasises executive management’s role in leading the implementation of data analytics (LaValle et al. ,2011; McAfee & Brynjolfsson, 2012, MIT Sloan, 2016). For instance, research show that when data goes against executives’ intuition, there is a great tendency to rely on "HiPPOS" - the Highest-Paid-Person's-Opinion. It hampers the development of data as a key resource in strategic decision-making. According to McAfee & Brynjolfsson (2012), "few things are more powerful for changing a decision-making culture than seeing a senior executive concede when data have disapproved a hunch." Many successful business leaders have built their careers on attributes like business intuition and industry experience. Having opinions challenged by analytics can create resentment towards further implementation (MIT Sloan, 2016). Management's support thus appear to be an integral resource in adopting machine learning.

Data scientists and analysts

Adopting and deploying machine learning also require people with the appropriate knowledge and skillsets. The traditional analyst's job in business intelligence has revolved around interpreting structured data, using historic figures. Interpreting big data, on the other hand, involves obtaining, extracting, manipulating and structuring sets of unstructured data, which requires a different skillset (Davenport & Patil, 2012). Davenport and Patil (2012) named data scientist "the sexiest job of the 21st century," describing the role as interdisciplinary and research-like. Because of the nascent state of advanced analytics and machine learning, data scientists in these fields advance by testing hypotheses and experimenting. Thus, the work process differs significantly from that of a traditional business analyst. Advanced analytics also address future concerns which require a deeper understanding of potential business challenges.

Machine learning requires advanced technical knowledge in mathematics, statistics and programming, combined with business knowledge. Not having the right talent can thwart attempts to adopt and deploy machine learning in analytics and thus companies must secure a critical mass of data scientist with the appropriate skills (Hoffman & Podgurski, 2013). Existing research address the interdisciplinary nature of the data scientist job, suggesting methods for adjusting academic curriculums to better suit the need of future employers (Jacobi et al., 2014).

4.6 Capabilities

Data Management Capabilities

Data management is a broad concept in IT management, but can be thought of as a collection of activities including (but not limited to) acquisition, recording, extraction, cleaning, integration, aggregation and representation of data (Gandomi & Haider, 2014). As data sets grow in size and become more complex, having the ability to manage data in an efficient manner becomes an important capability for organisations. In the process of adopting and deploying machine learning, data management capabilities provide a solid foundation from which more advanced capabilities in predictive modelling and advanced analytics can grow (pre-study interview).

The capability to connect data to strategy

Existing literature concludes that top-performing organisations "make decisions based on rigorous analysis at more than double the rate of lower performing organisation" and in high performing organisations, analytic insight is used to "guide bot future strategies and day-to-day operations" (La Valle et al., 2011; Sharma, Mithas & Kankanhalli; 2014). The benefits of analytics on business performance is evident, but for organisations to compete on analytics, data and analytics need to be integrated into strategic-decision making (Davenport, 2006). Organisations that have the capability to integrate data in strategic decisions, are likely to be more interested in analytical methods that have the potential to improve predictions and anticipate outcomes (pre-study interview). Thus, the capability of integrating data in strategic decision making will encourage the adoption and deployment of machine learning techniques in data analytics.

Educational capabilities
The capability to educate decision-makers about the potential and application of machine learning appears to be particularly significant for its adoption and deployment in organisations. It has been noted that machine learning algorithms are perceived as an opaque decision–making tool, which instils a level of mistrust in its outputs (Armstrong, 2015). For the likes of executives or business managers, it is important to have clear justifications for a decision. It is not good enough to rely on the supposed quality of the algorithm. This is particularly important when systems may be prone to errors or the decisions behind the choice of model is unknown. People understandably place more trust in humans than in machines, but the reluctance to trust new learning systems is a big challenge in realising their full potential. The capability to educate, inform and provide an understanding of how these systems actually operate has the potential to alleviate some of these trust barriers.

The capability to educate management in the potential and workings of machine learning can also get more managers to accept the value of data analytics. Introducing machine learning and advanced analytics inevitably shifts some power from employees with traditional expertise to data scientists and analytics (Galbraith, 2014). Many manual business processes can also be automated with machine learning. This shift in power from functional roles to data scientists and automated processes can make established employees feel less valued, and that their expertise is being questioned (Galbraith, 2014; LaValle et alt., 2011; MIT Sloan, 2016).This scepticism among managers halters the adoption and deployment of machine learning. Further knowledge among decision-makers on how machine learning and human expertise can benefit from each other to generate competitive advantage is thus needed.

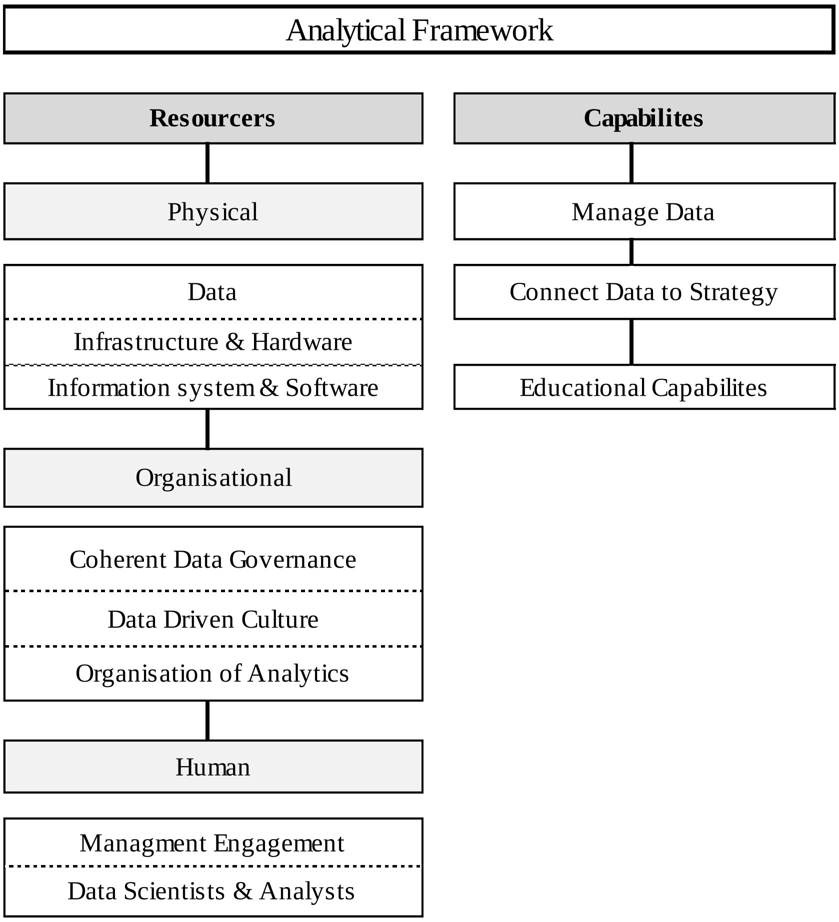

Fig. 10 Analytical framework of resources and capabilities

Chapter 5. Methodology

This chapter motivates the choice of a qualitative research method to study the organisational and managerial gaps that delay the adoption and deployment of machine learning in some organisations. It presents how data was collected, the interview design and how the results were analysed. The chapter ends with a discussion on the quality of the study.

5.1 Methodological fit

There is limited academic research on the adoption and deployment of machine learning from an organisational perspective. In order to explore what gaps in organisations are slowing down the progress of adoption in Sweden, an explorative qualitative research approach was chosen. Qualitative research is useful if the existing research on the subject is limited, and there is uncertainty about influential factors (Silverman, 2016). The study builds on existing research in related fields to narrow down organisational and managerial gaps with potential to influence the adoption and deployment of machine learning. Two frameworks were constructed to answer each research question. Using the systematic combining approach, the usefulness of the constructed models are tried on real cases (Dubois & Gadde, 2013). Semi-structured interviews were conducted with industry professionals in analytics and data science. The flexibility of the semi-structured interviews allowed different elements of the models to be explored in the different cases and facilitated an understanding for the social-constructs that underpins the identified gaps.

Research approach

The chosen research approach, which builds on abductive logic, was systematic combining (Dubois & Gadde, 2013). Abductive reasoning strives to explain a phenomenon based on the most likely interference that can be made from a set of observations (Flick, 2014). Systematic combining is a form of abductive approach where theoretical framework, empirical fieldwork and case analysis evolve simultaneously (Dubois & Gadde, 2013). The advantage of an abductive approach, compared to inductive or deductive reasoning, is the press it puts on the researcher to constantly reflect on the research process and challenge assumptions (Flick, 2014). It allows the researcher to freely move between empirics and theory, to incorporate relevant findings and elements in the theoretical framework along the way.

5.2 Data collection

Pre-study

In qualitative research, a pre-study can be useful to further the authors understanding of the subject field, especially if the existing body of empirical work is limited (Flick, 2014). For this study, a pre-study was conducted to explore potential perspectives to cover in a main study. It included three in-depth interviews with a sales director at a large Swedish telecom company, an experienced business and IT consultant with a background in machine learning, and a solutions manager at large technology corporation, responsible for selling machine learning solutions to companies in the Swedish market.

Interview sample

The nascent state of the adoption and deployment of machine learning in data analytics in Sweden made it difficult to study the process in-depth in a single case study. Upon the advice of two industry experts, who confirmed that finding relevant case studies in Sweden would be difficult, the decision was made to study adoption and deployment of machine learning in analytics across industries. It allows the identification of general gaps that are essential for organisations to bridge to enable the adoption of machine learning regardless of industry.

The decision was made to take on the perspective of those at the forefront of digital development: the data scientists and data analysts. This choice was made because previous reports (McKinsey, 2017; Servicenow, 2017) have approached the subject from a managerial perspective, surveying top management's opinions on AI and machine learning. The result of the pre-study nevertheless indicate an on-going strife between data scientists and business management, which suggested that many data scientist disagree with their management on the adoption and deployment of machine learning. To further investigate this potential conflict, and contribute with a new perspective on the subject, the view of the data scientists became the focus of this study.

In total 20 main interviews were conducted with individuals from 16 different organisations. The interview sessions lasted between 30-80 minutes in person or over the phone. In addition, two interviewees requested questions in writing and submitted written responses. To complement the main interviews, three follow-up discussions were conducted for clarification or to get more in-depth information on topics discussed. In the results, one interview was removed because it did not contribute with enough insight on the studied issue (see appendix 2 for interviewee overview and appendix 4 for interview guide).

The most important aspect in the search for interviewees was to establish their relationship to machine learning and analytics. Most had educational or professional training in data science, but a few interviewees had other backgrounds, but a strong interest in adopting machine learning within their organisation. Potential interviewees were screened through LinkedIn and also asked about their background. Interviewees furthermore asked for anonymity to share insights more freely and protect their respective organisations. It was granted as not specifying the specific individual or firm interviewed was considered to have limited impact on the results.

The aim of the study was to research the phenomenon across different industries, thus a choice had to be made regarding what industries to approach. There was a desire to cover a selection of industries from the top, middle and bottom of the AI index scale. Samples were then determined by how far they had come in adopting machine learning and to what extent they were allowed to share information. Some particular firms required non-disclosure agreements and extra paper work on the interviewee's part which made them difficult to access.

Interview design

Throughout the interview process a semi-structured interview approach, known as "responsive-interviewing," was used (Rubin & Rubin, 2012). The format allows the interviewer to ask open ended questions, relevant follow up questions and tap further into specific experiences of the interviewee, while maintaining a flow in the conversations (Rubin & Rubin, 2012). This conversational approach worked particularly well when more delicate topics were brought up. Some sensitive topics were discussed, and building good rapport with the interviewee was important for them to open up (Leech, 2002). Care was also taken to avoid asking presuming questions, stay aware of biases that may occur and demonstrate personal reflexivity to minimise the impact of the researcher on the data collected (Spencer et al., 2003)

Each interview required preparation and design of a general interview guide (see appendix 4). The guide was developed and refined as the research progressed. Gaining a deeper insight into the subject allowed questions to be refined and focus on topics that emerged as more relevant. Learnings from one interview could thus be used to dwell further into certain areas in the next. Research on the industry, firm and interviewee was done before first contact was made. If an interview was booked, more detailed research on the firm and individual was done, relying foremost on news articles, educational online videos and LinkedIn.

Data processing

Interpreting data is described as the core of qualitative research (Flick, 2014). The interview data was interpreted and sorted based on common overall themes. Quotes were extracted from the interview transcripts and categorised in a spreadsheet to detect commonalties and differences. Categories were determined with the help of the theoretical framework, and findings were consistently evaluated against the theoretical background, making use of the abductive research approach.

5.3 Quality of Study

Credibility

Credibility refers to the integrity of the results and whether the representation of data epitomise the views of the studied participants (Noyes et al., 2011). Ensuring credibility is one of the most important aspects to establish trustworthiness in qualitative results (Lincoln & Guba, 2000). In qualitative research, credibility is highly dependent on the ability and efforts of the researcher. In this study the author took precautions to secure credibility by (1) adopting a well-established research method, (2) through examination of previous research findings in adjacent fields, such as big data analytics, to asses that result are congruent with past findings, (3) match findings from people in organisations with those of industry experts to see if results concur, and (4) agreeing to anonymize participants to enable them to speak freely and honestly with no concern for repercussions.

Dependability

Dependability refers to consistency and reliability in qualitative research (Shenton, 2004). If the project was repeated with all the defining parameters staying the same, would it generate the same result? This can be problematic to prove when researching a dynamic phenomenon in a changing environment. According to Lincoln & Guba (2000), dependability is closely related to credibility, thus ensuring credibility also generates a degree of dependability. To address the dependability of this study the author has given a detailed account of how data was collected, and evaluated the effectiveness of the chosen methods (Shenton, 2004). Dependability in this study is limited by anonymizing the studied organisations and interviewees.

Transferability

According to Merriam (1998) transferability refers to the extent the results of one study is applicable in another situation. There are diverging views on the transferability of qualitative research. Erlandson et al. (1993), argue that because qualitative studies are concerned with a small and specific sample, the findings are non-transferable because they are defined by the context of the original study. However, other researchers disagree (Denscombe, 1998), claiming that even a specific case is part of a larger group and certain aspects of the findings are thus transferable. To aid any transferability of the findings in this study, the author has taken care to convey the boundaries of the data collected as well as the results. Through a multiple-case design and purposive sampling, the author has also made an effort to research the same phenomena in different industries, to increase transferability of results to different business contexts.

Chapter 6. Empirics

The following chapter presents the results from the in-depth interviews with data scientists and analysts in 16 different organisations which are adopting machine learning in big data analytics. First, this section presents the identified organisational and managerial gaps impeding organisations' adoption of machine learning. Second, essential resources are presented. Third, two examples are presented of how organisations have combined resources to build new capabilities in order to bridge identified gaps.

6.1 Identified Gaps

The findings reveal gaps in following attributes which impede the adoption of machine learning:

Adaptability

A fundamental challenge for many of the organisations studied is their long legacy. They carry a substantial heritage of organisational hierarchies, practices, routines, culture, competencies, technology and more, which set the context for new developments. How a firm operates today is a result of its legacy and there is a noted gap between the current state of operation and what is desirable for implementing machine learning.

Traditional project management routines

Several analysts note that established routines, demand on results and need for benchmarking are example of routines and practices slowing down the adoption of machine learning. A customer developer reflects over how the firm's set routines for initiating new projects are making it difficult get an AI or machine learning project going:

"…all new projects in the firm go through the same process: first you present the project plan and motivate how this investment will generate business value. Managers want to know the timeframe, specific results, ROI, and how it affects the bottom line. It is impossible to know those things when implementing a machine learning project for the first time. This rigid form for initiating projects are thwarting our attempts to develop AI and machine learning..." (#10, see appendix 2).

Other interviewees agree that innovating around machine learning projects often challenges management's demand for timely results. Another data scientist explain how their machine learning project started with sorting out the company's databases. It took them three years before they had the necessary data warehousing and infrastructure in place to start experimenting with machine learning. Until this day the project is yet to prove profitable as the firm has only done pilot projects. These type of projects depend on management's confidence in that they will eventually generate business value for the firm. However, the time it takes to develop machine learning solutions and the difficulty to estimate return on investments are noted to give executives cold feet, and make them reluctant to prioritize those type of investments.

Another mentioned difficulty is benchmarking the outcome of many machine learning projects. Traditionally, firms are concerned with benchmarking and evaluating results. What constitutes a good result in machine learning and what are reasonable expectations on the outcome? One data scientist explain:

"…at university they teach you what a good algorithm is, but those measurements mean nothing to business executives. We are speaking different languages, which makes it challenging to communicate around these issues and manage their expectations…" (#8).

The interviewees explain that what constitutes a good result in their eyes may not meet management's expectations. For instance, a data scientist working with natural language processing often find that people are comparing the results of algorithms with the sophistication of human language. What constitutes a good result from a machine learning perspective, may not seem that good in the eyes of a layman. Many note being more lenient on routines for initiating and evaluating machine learning would benefit adoption.

Cultural heritage

Another identified gap relating to a firm's legacy is aspects of the company culture. The adoption of machine learning benefits from a company culture that welcomes new technology and is data oriented. Some interviewees nevertheless note that their organisational culture is far removed from technology and lack a general interest in analytics. The gap makes it challenging to adopt big data analytics and machine learning. A head of analytics explains:

"…at the heart of our business culture, we are a sales organisation. We still have people walking door-to-door selling our products. People in the organisation have worked their way up and many don't have an academic background. They are removed from technology and many are unable to see the potential it holds. Some even feel threatened by it, and are actively working against new developments because they prefer the old systems and routines…" (#9).

This sort of cultural legacy is a significant road block in moving forward with machine learning initiatives. There are also internal power structures to consider. Some interviewees acknowledge that the shift in power away from some traditional roles are not appreciated. A solution manager explains:

"…it is important to remember that not everybody benefits from analytics. For instance, in our firm the product development department carry a long legacy and is very well respected. They have a very strong influence in the organisation, marketing and sales just follow suite. Let's just say they are less than thrilled about having a data scientist or analyst tell them what the next big trend is and what they should be developing. These people consider themselves "masters of their art"…" (#3).

The statement depicts how difficult to it can be to address these power shifts and create an interest for adopting machine learning. On the other hand, another head of analytics reflects on how carrying less legacy has benefited the integration of data and analytics, even in more creative professions:

"…I anticipated that it would be difficult to get our creatives and game designers to care about analytics, but they have been surprisingly good about it. They are interested in knowing how the customer behaves and get insights on what aspects of the games are working and not…. I think it's an advantage that we are a relatively young organisation with younger people. People are not as set in their ways, there's less prestige and the culture is quite open-minded…." (#13).

It demonstrates how cultural legacy is affecting the adoption of advanced analytics and to successfully implement machine learning, the drawbacks of cultural legacy must be bridged.

Technical heritage

Several interviewees also refer to their technical heritage, or "technical debt," as a challenge. There is a substantial gap between what many of the organisations' aged IT infrastructures can do, and what is required for using new data sources, implementing machine learning solutions, applying new software etc. An analyst explains that part of their information systems have been in place for over 40 years, and bringing those systems up to speed to work with the variety and volume of new data takes time and require new investments.

Up-dating the infrastructure is important to make the most of new data sources, and connect them to the system. A head of analytics explains:

"…I see a scenario where we, in two-three years, are using sensor data from our project and analyse data in real-time. However that requires an up-grade of our infrastructure and our capacity to handle data fast....we talk about time-to-market: how long does it take from the point data is collected, to analyse it and generate insights that create value... to deliver fast results we need an infrastructure that can keep up…" (#16).

Although bridging the technical heritage gap may seem more straightforward in comparison to the gaps in routines and culture, interviewees note that prompting investments in IT can be difficult. IT systems are considered support functions and there is a tendency to favour investments in projects where there is a more direct connection to driving business value.

Efficient Organisational Set Up

The findings reveal that several of the organisations are experiencing challenges with the organisation of their analytics competence. The organisation of analysts and data scientists within the organisation is described as important in order to link innovation projects to relevant business practices. In addition to how the competency is organised, firms must consider where in the organisation the new competency should be located.

The most common form of organisation among the studied organisations is the centralisation of analytics competency. For some companies, centralising know-how is the only way to get a critical mass of people with enough skill to successfully innovate around new knowledge areas such as machine learning. The major issue with centralising competency is nevertheless that it creates a distance between the people who have the know-how and the people who have the problem. One customer developer explains:

"...the knowledge in AI and machine learning sits in the IT department, but they don't have a need. I see how new tools could improve practices in my area, but I don't have the competency. There is no natural way for us to interact or exchange ideas, our departments are even located on different floors…" (#10)

The challenges are echoed by others, who reflect on the difficulty of knowing where in the organisation to innovate around AI and machine learning. It may seem natural that the competency should sit in the IT-department where most data scientists have belonged in the past. However, several interviewees note that this is not the ideal place as the competency becomes too distant from core business practices. IT is generally considered a support function and if the goal is that new solutions in machine learning should drive business value, the competency has to be closer to business operations.

One organisation is also experiencing drawbacks of having their advanced analytics competency located one business area. Their customers are nonetheless not limited to one business area and to study customer behaviour, data is collected across functions. It has resulted in internal confusion regarding the role of this new competency centre and how it fits in with existing business functions.

Other organisations have attempted a decentralised model, and work in cross-functional teams. The benefit is that the competency is close to business practices, but it makes innovation more fragmented. Different competencies evolve in different teams. The major drawback of decentralisation is described as "...the fact that our data scientists never get to benefit from each other's work as they are located in different business units…". Innovating in machine learning involves a lot of trial and error. Hence, learning and from the wins and losses of others and exchange experiences are key to gaining momentum in developments. When data scientists are located in different business units there is no natural way for them to interact or collaborate, thus they rarely end up sharing experiences.

An additional problem with decentralisation is the up-keeping of a common data governance system. Interviewees explain how data scientists in different business areas came up with their own standards, methods and benchmarks tailored to the needs of their team. Although there are noted benefits with cross-functional teams, others indicate that organisational barriers can hinder innovation. One data scientist note that it can be difficult to get the financing and support to innovate around machine learning if you are in a business unit with their own agenda, budget and goals. Supporting the advancements of AI tools is rarely the top priority, which halts the adoption of machine learning.

Data-driven culture

Analytics and the adoption of machine learning benefit from a data-driven culture, but many organisations lack a data-driven decision culture, creating a significant gap. Some interviewees describe their organisation as data-driven in short-term or operational decisions. It is less common to let data drive strategic decisions. There are several mentions of prominent tech giants, and everyone admits that their corporate culture still has some way to go in comparison. There is widespread use of descriptive data, follow-up of KPIs and reports on historical figures. There is little or no integration of data in strategic decision making at all levels of the organisations.

When asked about the reasons for not being more data-driven, there are a few fundamental factors that stand out. Again, legacy is a considered a major reason. At executive level there is the tradition of relying on intuition and experience. Data has a supporting function, rather than a leading function and is generally used to confirm decisions rather than drive them. Data and analytics can be an inconvenience as well. Some note that a more data-driven culture is being hampered by people who wish to hold on to their "visions" or "intuition". A consultant describes his experience of working with advertising agencies in Sweden:

"...take marketing for example. If you look at advertising agencies in the US and England the majority of their staff are media planners, strategists, analysts. They have one or two creatives. In Sweden it is the other way around. Here marketing is about implementing a creative vision, rather than relying on data to drive sales…" (#19).

Naturally, others disagree with this statement. There is nevertheless an identified difference in culture between business areas. Functions with operations online such as online marketing or sales, were said to be more data-driven because it has been part of their "culture" since these business areas were introduced. It was noted that for advanced analytics and machine learning to be of any value to the business, there needs to be a widespread change in the mind-set of the entire organisation regarding data and analytics.

Management's Knowledge

Interviewees note that, as data scientists and analysts, it can be difficult to discuss technological advancements with stakeholders in the organisation because they lack an understanding of what it is and how it works. An essential challenge in adopting machine learning is bridging the gap between those who have the knowledge (data scientists and analysts) and those who doesn't (stakeholders in the organisation). This gap in knowledge is described to underpin several of other reasons to why the adoption of machine learning is slow, such as a lack of trust in machine learning methods, a lack of support from management, reluctance to invest in machine learning-projects, unreasonable expectations of results etc.

There is an apparent frustration with management's inadequate understanding of the implementation process for machine learning among the interviewees. Some explain how they have tried to implement AI project for years, but management's reluctance to put forth the money to see them through thwarted any developments. The reason for management's hesitation were thought to be their failure to understand how advancements of machine learning could generate business value in the future. When compared to other possible investments, the exact returns of AI is unsure and thus those projects appear more risky.

Among interviewees in organisations where management is supportive of AI and machine learning, there is still a frustration about management's poor understanding for the implementation process. Managements are willing, and in some cases, eager to move ahead with data analytics and machine learning projects but they do not grasp the process. A head of analytics explains:

"...because management doesn't know how machine learning works, they don't understand how to implement it and they are unable to articulate what needs to be done. You can't simply hire a few data scientists and have them come up with models, implement them and expect revolutionising results. There is very little understanding for the fact it takes time to get an organisation ready for adopting machine learning and apply advanced analytics… and I'm the one who has to educate them" (#9).

Another data scientists elaborates:

"....I'm a data scientist, not a miracle worker. Management's anticipation on what I was expected to achieve was unreasonable, especially within that timeframe. They did not understand the amount of groundwork that have to be done, before it is possible to design analytical models and apply machine learning…" (#5).

Closing the knowledge gap between data scientists and managements is essential to enable organisations with the preparation work that has to be done before it is possible to adopt machine learning solutions.

The Organisation's Knowledge

The findings also indicate that simply educating top management is not enough. There are cases where management is advocating for AI initiatives but the middle management is not on board. Similar issues arise when other functions of the organisation feel threatened by new developments or are unwilling to support new technology advancements. Consequently, there is a significant gap in knowledge about analytics and machine learning between data scientists and employees in the rest of the organisation that has to be addressed.

In order to realise the full potential of implemented analytics tools and competence, it is essential that there is a universal understanding for how to utilize these resources. A head of analytics explains:

"...for example, sales are concerned with following up last week's figures, they don't ask our analysts 'tell me what I can do to increase sales for next month' because they don't know that's the sort of question we could answer with predictive analytics and machine learning. It is not enough that our data scientists understand predictive analytics, our sales people must have a grasp of it too..." (#17).

Another head of analytics develops:

"…we've just gone through a reorganisation so that my team [the analytics team] will get more resources to work with AI and ML. But to most people these are just buzzwords that they've heard you are supposed to work with. It is interesting how nobody (except for the people on my team) has a clue what it is, what to do with it or how to implement it..." (#19)

The situation illustrates the demand for an increased level of awareness on advanced analytics and machine learning across the organisation. Some analysts refer to it as increasing the "analytics maturity" of the firm. It indicates moving away from reporting on historical data and start using predictive modelling, implementing visualisation tools for business units to have direct access to data, and develop a culture where analysts and data scientists are working with data to solve strategic problems. An analyst explains:

"…I would say there are four parts to data analytics: processes, competence, technology and data. They are all integrated and to increase the firm's analytic maturity, these four parts have to be improved simultaneously. It is not enough to simply invest in one or two areas… our management is eager to move ahead with machine learning, but for our organisation it is a big step to move from working with traditional BI to visualisation. You can't skip ahead. Even though we invest data and technology, we don't yet have the processes and competence…" (#16).

It demonstrates the need to increase knowledge across the organisation and approach the implementation of machine learning as a process that evolves in stages, rather than something that can be acquired and implemented straight away.

A lack of knowledge also results in a lack of trust for new tools and technology. Data scientists describe that a significant part of their job is building people's trust in the technology. Implementing machine learning can for instance automate processes in sales and marketing. It nevertheless means trusting a machine to make decisions and carry out tasks, and it can be difficult to gain that confidence from someone who does not have a fundamental understanding of how the technology functions.

Talent

There is a general agreement that acquiring the right competence and knowledge is vital for adopting and deploying machine learning. There is an identified gap between the talent organisations currently have and the talent they need to adopt big data and analytics. Some have for instance never employed data scientists before, but simply outsourced the job. A majority of the organisations are currently recruiting data scientists with machine learning competence. It is a challenge for the organisations to find the right people – and keeping them.

Most organisation state that "the right people" have a strong skillset in programming, mathematics and statistics in combination with business experience. Because many of the organisation are at the outset of developing solutions using machine learning, there is furthermore a need for talent who enjoys experimenting and researching possible applications.

Organisations are also starting to place more importance on personality and the ability to communicate around analytics. An analyst explains his dilemma:

"….I have met with data scientist that are great at programming and building models. But today a data scientist's job also requires great interpersonal skills. Especially as data and analytics become an integral part of our business and organisation. Data scientists need to be great communicators to initiate collaborations and interact with people across the organisation. Finding the individuals who are outstanding in both areas is challenging…" (#15).

In addition to finding and requiting the necessary talent, several department heads also agree that keeping and developing talent are the greatest challenges of their job. Experienced data scientists, who also understand the business perspective, are highly sought after. Several department heads explain how they work hard to retain their talent as they are offered new opportunities almost every day.

It is, however, noted that some firms are at risk of losing talented data scientists because their abilities out-grow the organisation. A consultant explains:

“...I have seen it happen in many organisations. There are talented individuals who are passionate and knowledgeable about data science and machine learning. They work on initiatives, but eventually they reach an “intellectual glass ceiling” when they don’t win support from managers or executives. Eventually they move on, and their passion projects die. It is a great loss for the organisation, but they don’t realise it… yet…” (#19).

Many of the interviewees are in similar positions themselves and are pushing for their organisation to adopt and deploy machine learning through “passion projects.” Bridging the gap and securing the right talent is essential for organisations to adopt machine learning.

Data Collection Processes

Big data analytics require great sophistication of how data is collected and treated internally. There is an identified gap between how data is collected and managed in firms today and the management and treatment that would facilitate the application of machine learning.

One fundamental challenge for many companies is the lack of a coherent system for collecting and storing data. As one head of IT comments:

“It sounds strange, but we don't even know what sort of data we have. For years and years databases have simply been filling up with data without anyone taking a closer look at it…” (#7).

The prevalent method of collecting and storing data is in separate information silos which contain years of aggregated structured and unstructured data. Data scientists are spending their time retrieving, organising and restructuring old systems to create the necessary structure to be able to work with the data. It is a time-consuming process, and some have spent years sorting out their databases, while other organisations have merely started. Big data analytics benefit from centralised storage of data that makes it possible to combine different data sets.

The underlying reason for the fragmented data systems is the lack of reason to alter it when posed against the time, effort and ability of doing it. Most firms in the study was in business long before the digital era. Their systems have developed successively as technology has advanced as more data could be collected.

A data scientist from the energy sector explains the widespread problem:

“...take companies like Spotify for example, it is easier for them to be data-driven in everything they do because collecting and analysing data have been a key activity for them since they started. It is different for us, we carry a different legacy and our systems were never designed with big data and machine learning as a possibility…” (#6).

Although many argue that younger firms benefit from evolving in a digital era and have adapted their data governance system accordingly, one data scientist disagrees as he argues that it is simply a matter of time before younger organisations will face similar challenges:

“...the challenges with data governance develops over time. There will come a time when younger firms will have to deal with their legacy and adjust their data governance routines. The benefit of being a young company does not last…” (#8).

The relative frugal use of data and analytics are noted as a reason to why organisations have not addressed weaknesses in their collection and management of data earlier on. Big data analytics and machine learning are nevertheless changing that.

Business Relevance

Big data analytics and machine learning are expected to generate great competitive advantage, but only if they drive business value. There is an identified gap between machine learning initiatives that drive business value, and the machine learning projects many firms are currently investing in. There is evident frustration among the interviewees about some organisations' inability link machine learning projects to business purposes. An analyst explains:

"...you get these 'innovation hubs' where a group of people play around with ideas and it gives management the chance to say 'we are investing in AI.' The problem is that these people are completely removed from the day-to-day business operations. No one is considering solutions with a specific business problem in mind, like how we can improve our customer satisfaction…" (#16).

It highlights a division in thought regarding how to innovate around machine learning. There are those who appreciate the "innovating for the purpose of innovation" approach, where data scientists and analysts freely innovate around AI, machine learning and analytics in "hubs." Others strongly advocate a more purposeful approach, where solutions are developed to improve a specific situation.

In one of the studied organisations both approaches were pursued, but by different teams. There was a noted scepticism towards the 'innovation hub' with the data scientist involved in the project which is essentially rebuilding the organisation's main platform, using machine learning. Although politely expressed, the apprehension appeared founded on the fact that output from the hub had diminishing relevance to the organisation's operations. The manager of the hub explained that the team do not have any specific demands to fulfil, nor are their results evaluated or benchmarked. The team in the hub is free to innovate around what interests them, and to experiment with new ideas.

Other interviewees echo the need for machine learning projects to contribute in achieving strategic goals and driving business profits, in order to see an upswing in interest. A data scientists explains:

".....many companies are boosting with their investment in AI, but you have to do more than implement a Chabot for your customer service. Yes it is 'a machine learning project' but it is highly unlikely to generate a competitive advantage for your business…" (#17).

In order for big data analytics and machine learning to constitute ad competitive advantage, it is important that initiatives have substantial business relevance.

Summary of Identified Gaps		
Identified Situation	**Gap**	**Desirable Situation**
Old routines, practices, standards, infrastructure etc. associated with the firm's heritage are holding back ML adoption	**Adapability**	New project management routines, open-minded corporate culture, improved infrastructure to enable the adoption of ML
Lack of data driven culture. Data is mostly used to support decisions. Data is not driving strategic decision-making.	**Culture**	A data driven culture where data and analytics are used to drive both long- and short-term decisions
Centralised or decentralised analytics teams which are either too removed from business practices or too fragmented	**Efficient Organisational Set Up**	A hybrid organisation of analtyics, integrating centralised and decentralised structures
Data is collected in sperate silos, preventing integration. Different formats, standards, labels. No common management system	**Data Collection Processes**	Data is collected and managed in a centralised system, using common standards
Traditional analysts, lack the competency of data scientists. Inexperienced data scientist with little exposure to business	**Talent**	Data scientists that are skilled in both data science and business, with great interpresonal skills
Managemen lack understanding for how to implement BDA/ML. No ability to articulate what needs to change	**Management's Knolwedge**	Management understands the implmentation process of BDA and ML
Few employees other than data scientists understands BDA and ML or know how, where and when to captilise on it	**Organisation's Knowledge**	There is a widspread general knolwedge about BDA/ML and employees across the organisation can identify opportunites to exploit the technology
ML projects are iniated for the sake of innovation. They have little business relevance.	**Business Relevance**	ML projects are initated with intent to drive business value. There is a plan of action

Fig. 11 Summary of gaps: the following gaps in organisational and managerial attributes have been identified to delay the adoption and deployment of machine learning in organisations (for an overview of identified gaps in each organisation see appendix 3).

6.2 Resources

This section presents the identified resources that are essential in adopting big data analytics and machine learning

Physical Resources

Data

Data is a significant resource for all of the studied firms and the fundamentals of advanced analytics and machine learning. The studied organisations have access to large quantities of data from all business functions. It is also common that organisations purchase external data.

Machine learning algorithms not only require large amounts of data, but the quality of data is essential to generate reliable insights. Quality refers to the completeness, accuracy, consistency, timeliness, granularity of data sets. In the creation of machine learning models, data scientists need training data, to train the algorithms. The quality and accuracy of training data is particularly important as using inadequate data can result in a poor model. Many note that accessibility to quality data is a fundamental challenge in adopting machine learning. The difficulty with data quality is often related to how data has been treated in the past, and to develop machine learning models historical data sets are needed as well. A data scientist explains:

"…if you look the data you get from Google Analytics today, then the quality is pretty good. But if the data you look at is a conversation a sales representative had with a customer three years ago, the quality is something else…" (#12).

Because of the varying quality, data scientists have to spend a lot of their time "cleaning" data. One manager estimates that the data scientists on his team spend up to 80 percent of their time improving the quality of data to make it usable for analytics. Because cleaning data takes a lot of time, data scientist note that it is highly important that functions across the organisations are educated in how to manage data. Implementing universal procedures for how to collect and treat data ensure better data quality, and allow data scientists to dedicate more of their time towards building analytical models rather than cleaning data.

It-architecture and hardware

Collecting and analysing big data require a certain level of server and computer processing power (because of security reasons, all data cannot be kept in the cloud). A manager from a younger firm recalls how the organisation has had to limit the amount and form of data they collect, because of their limited storage capacity and infrastructure. The increase of data generated from mobile devices has nevertheless resulted in the need to improve storage and process capacity. Others also refer to the increase in data from mobile devices as a trigger to increase the capacity of computer resources. Machine learning furthermore requires robust systems that can handle a large flow of continuous data (unlike traditional data which come in batches).

Information System and Software

In big data analytics and machine learning also involve the use of new software and data scientists use a lot of open source libraries. Apache Hadoop is an open-source software framework for storage and processing of big data, which is used by several organisations. In the construction of machine learning solutions open source software libraries such as Google's TensorFlow and Keras are used. These open source libraries make it fairly easy to experiment with different types of machine learning, such as deep learning and neural networks. Several interviewees mention how they even spend their free-time studying new software to experiment with machine learning.

Organisational Resources

Management Support

The support, or lack thereof, from top management was widely discussed among the participants. It stands-out as a significant resource for AI projects and machine learning initiatives to move ahead. Having top management's approval unlocks financial resources and gives a different mandate to operate within the organisation. Everyone agrees that having top management's support greatly facilitates the adoption of advanced analytics, machine learning and other AI technology. A Head of Analytics recalls how pivotal the collaboration with the company's CIO have been in achieving the necessary organisational change to integrate analytics into business practices, and take steps towards adopting machine learning to study customer behaviour. Another data scientists explains how the introduction of a new IT-manager had profound impact on the innovative culture in the department. The new manager introduced a less hierarchical work process, and provided great support for introducing new initiatives. It resulted in the introduction of machine learning projects, which are still in progress.

For those who have had to promote data analytics and the adoption of machine learning with limited support from their management, indicate that they have had to spend a significant amount of time "selling" their ideas to the organisation. Many are driven by a strong interest in AI and have taken the time to read-up on the technological developments through online resources and refined their skills in machine learning using open-source libraries. The motivation is described to come from personal curiosity, but also from seeing the business value the new technology can add. It is apparent that without the passion of these individuals, the organisations without management's support for AI would soon be falling behind in the adoption of machine learning.

Business Units

Business units refer to the other functions in the organisations. They are both the users and producers of data. Data scientists and analyst serve business units with data and analytics to facilitate their decision-making process. It is easy to focus on the importance of data scientists to adopt machine learning but in order develop advanced analytics capabilities, the business functions must keep up with the progress as well. Data scientists, especially if sat in one team, do not have the same insight into operational businesses challenges as the staff in the business units. Thus the knowledge and experience of the business units is essential in developing the understanding of how machine learning can be applied to address pressing business challenges. An analytics managers explains:

"…initially product owners had very poor knowledge about the sort of questions we [the analytics team] can help them with and but it is a learning process that develops through interaction and dialogue. It is also difficult for a new data scientist to know how to generate value for the product owners as well, especially if they are new to gaming – it goes both ways…" (#13).

Human Resources

Analysts and Data Scientists

The analytics team generally consist of traditional business intelligence analysts, which are skilled in statistics and tools for traditional process optimisation and forecasting. Big data analytics and machine learning techniques foremost require knowledge in programming and mathematics. As previously mentioned, the new growing role of analytics in organisations increases the demand for data scientists who are business savvy and excellent communicators. If data scientists are natural communicators and presenters, it will facilitate the spread of both explicit and tacit knowledge throughout the organisation.

The different titles can lead to some confusion as ‘data analysts’ or ‘heads of analytics’ can refer to someone with either a traditional background in BI-analytics or a data scientist.

Analytics Managers
Many of the interviewees are department heads or director of analytics and fairly new in the roles. When asked about their job, several start off with “…I was hired to develop the analytics team…”. These individuals have been recruited for their strong background in data science and advanced analytics, along with the conviction that they are able to spread this knowledge to the rest of the organisation. These managers are “analytics advocators” and carry a significant responsibility to education and inform the rest of the organisation about the benefits and potential of advanced analytics and machine learning. One of the most valuable assets these people bring is their ability to engage in pioneering actives through their understanding of how resources can be combined to form new capabilities. They also have the theoretical and practical knowledge to think creatively about or how their knowledge branch relate to other knowledge branches. A director of a health-care provider explains:

“…this project happened by chance. I’m cardiologist with a background in business and a strong interest in digital developments. At a conference I met a logistics manager for an automotive company who told me how they apply machine learning to optimise logistical processes and predict interruptions that would be costly for their systems. I realised that same technology would be helpful in health services. Patients suffering from heart failure are the most expensive for our system. Through this new initiative we use machine learning to predict the results of their treatments, the likelihood of them being readmitted or dying” (#18).

The example illustrate the value of having a basic understanding of machine learning and analytics in order to identify new areas of application. Analytics managers are hired to do it, but if more individuals in the organisation have a fundamental understanding of the technology they can contribute with these type of valuable insights into their areas of expertise as well.

6.3 Capabilities

This section describes two examples of how organisations have combined resources and developed new capabilities to bridge a few of the identified gaps

Organising competency
To address the challenge of organising analytics competence, one of the studied companies have come up with a hybrid model. By combining human resources (analytics staff and business units) in a unique way, the firm has built the capability to capture the benefits of both the centralised and decentralised model. The organisation’s data scientists spend 70 percent of their time working together in a centralised analytics team. It allows to them to benefit from each other’s know-how and innovate around machine learning from an enterprise perspective. The other 30 percent of their time, the data scientists spend with different business units. It gives them the opportunity to stay in touch with the every-day business challenges and provide them with a deeper insight into the issues their co-workers need their help solving. So far, the organisation is pleased with how this hybrid model is working for them.

Educating the Organisation & Becoming more data driven
A couple of the studied organisations have managed to develop capabilities to narrow the knowledge gap in their organisations and become more data driven. They have implemented visualisation tools that visualises data in a logic and pedagogical which makes data more available to other functions in the organisation. An analytics manager explains:

"…we are implementing a new BI tool to make data more available to the decision makers in the organisation. They can then access the data they need which will be valuable for them… I think this will be the key to making our organisation more data driven…" (#13).

Visualisation makes data and analytics available to business unites and management. It levels out the knowledge gap in the sense that it allows the organisation and decision makers to become more familiarised with the data available to them. Visualisation tools contribute to making organisations more data driven as they make data more readily available and facilitate the integration of analytics in their every-day decisions.

Chapter 7. Analysis

This chapter starts off by identifying capabilities organisations can develop in order to bridge the identified gaps. The second part of the chapter discusses what resources and capabilities organisations should develop to facilitate the adoption of machine learning.

Part I: The first research question states: *What are the gaps in vision, knowledge creation and knowledge integration/application that impede the adoption of machine learning*?

The section below outlines how the gaps in the identified organisational and managerial attributes result in gaps in the organisation's vision, knowledge creation and knowledge integration/application.

7.1 Gaps in Vision, Knowledge Development and Integration/application

Vision

Recognising machine learning in long-term strategies

The findings demonstrate that there are diverging views on including machine learning in the organisation's future vision. The lack of management support that some organisations are experiencing is reflective of a vision gap, where the organisation's executives not yet envision machine learning as part of their organisation's long-term strategies. In other organisations management is supportive of adopting machine learning but lack the knowledge to strategise on the subject. Machine learning is in their strategic view field, but they are unable to come up with a viable plan for how to realise it. However, the fact that almost all of the organisations are recruiting talent in machine learning show that they are investing in their vision. It seems as though the requirement of department heads and middle managers in analytics play an important role in anchoring management's vision about machine learning in more tangible strategies and activities. On the other hand, findings reveal that some organisations are underinvesting in IT-infrastructure, suggesting that they may still be subject to "bounded vision." As the findings demonstrate, capabilities in machine learning depend on several factors, only investing in talent is not enough.

Knowledge Creation

Effective learning system

The studied organisations lacked a highly systematic way for sharing knowledge. As the identified knowledge gaps suggest, this has negative consequences for the application and adoption of new technology. A few organisations have experienced positive results from including formal training sessions, but there is no long-term plan for how to improve knowledge or increase learning in the firm. It is a significant gap that has repercussions for the innovation process. For a further discussion on how to bridge this gap, see the section on education capabilities below.

Effective exchange of knowledge

(1) **Data-scientists to data scientists**: The findings illustrates that organisations struggle with how to organise their analytics competency – in a centralised team or embedded in business units. There is an identified gap in exchange of knowledge between data scientists in the decentralised structure. Keeping the data scientists in business units increased the effective exchange of knowledge to the organisation, but it removed a natural way for data scientists to interact with

fellow data scientists in other teams. The lack of socialising among data scientists result in limited exchange of tacit knowledge which according to theory has a negative impact on innovation. In was mentioned as a problem by the interviewees that data scientists become too distant from each other they cannot gain an advantage from each other's experience (tacit knowledge), elongating innovation processes. To bridge this gap, the hybrid model as implemented by one organisation, appears to be effective.

(2) **Data-scientists to management/organisation:** Referring back to the knowledge gaps, it is evident that there is a gap in the exchange of knowledge between data scientists and management/organisation. In addition to training, an example to address this gap is the use of visualisation. By developing the capability to visualise data and spread available of data to management and business units, it can level out one aspect of the knowledge gap. For further discussion on how to bridge this gap see the section on education below.

Knowledge Integration and Application Gap

Ability to integrate data tacit knowledge

Tacit knowledge is described as a significant resource in innovation. Theory holds that tacit knowledge is spread through socialisation, thus it is essential that organisations provide natural points of interaction between data scientists and the business units. The findings reveal that when the competency in data science is centralised or placed in hubs, it greatly hinder the integration of tacit knowledge between data scientists and the business units. It hinders the adoption of machine learning because people in the business units have problems that the technology could solve, but there is no exchange of ideas or knowledge because the people involved are not naturally interacting. The findings show that some organisations have tried implementing "hackatons" and "innovation days" with the purpose of stimulating the exchange and integration of tacit knowledge. The problem is that these theme days are often designed for data scientists, thus limiting integration to a small group. Moreover, integration requires organisational routines that allow continuous exchange of tacit knowledge over time. The hybrid model for organising analytics competence, where data scientists move between a centralised team and the business units, is a good example. The organisation has managed to 'routinize' interaction, through implementing a specific work routine which allow a systematic integration of tacit knowledge. For further discussion on how to bridge this gap see the section on education below.

Part II: The second research question asked: *What resources and capabilities should organisations develop to bridge identified gaps?*

The following section outlines the essential resources and provides examples of capabilities that can be built to bridge some of the identified gaps.

7.2 Essential Resources

This study aims to research the essential resources to adopt and deploy machine learning in analytics. The findings reveal seven key resources for adopting machine learning in analytics: *data*, *IT-architecture, software*, *management support*, *business units*, *analytics staff* and *analytics managers*. Of the physical resources, data and IT-architecture stood out as particularly important because of their fundamental role in big data analytics and machine learning. Combining data and IT-architecture with other resources allow companies to create significant capabilities to adopt and deploy machine learning.

Management support was noted as an essential organisational resource. Having the support of management proved important because of the many other resources they control. As findings disclosed, building new capabilities without management support limits access to financial resources and time resources for instance. There is no question that management support is essential, but it could be one of the most difficult resources to acquire. How do you win management's support for adopting machine learning projects?

The empirical findings reveal that data scientists believe that demonstrating experiments that can prove positive results on a small scale is the best way to win management's interest. Another suggestion would be to, to the extent it is possible, use A/B-testing to prove in real numbers, how the impact of a specific initiative are affecting the customer-base for instance.

In difference to the analytical framework, findings also disclosed business units as an essential resource. Despite previous research's emphasis on management's role, business units are a key resource in the adoption of machine learning. As demonstrated by the findings, more employees than the data scientists must be apt in applying analytics and locating opportunities for machine learning in order to fully capitalize on the technology's potential. The business units are an essential resource in building those capabilities.

7.3 Essential Capabilities - Bridging the Gaps

The following section discuss how organisations can combine available resources to develop new capabilities to bridge the identified gaps. The most essential capabilities are identified: *connecting analytics to strategy* and *educating management and the organisation*.

Connecting Analytics to Strategy – Bridging the Gap on Business Relevance

A gap related to the business relevance of machine learning projects was identified in the study. Big data analytics and machine learning can generate significant competitive advance but several machine learning projects in the studied organisations lacked direct business relevance. A capability organisations can develop to address this gap is connecting analytics to strategy.

Connecting analytics to strategy require the combination of several essential resources – data, analytics staff and management support. Referring back to RBV, connecting analytics to strategy can be thought of as an enriching activity. Most organisations already have a certain level of analytics capability, but to drive business value, it needs to be extended to include strategy. There are several actions organisations could take to build this capability.

In a first step organisation should consider where the analytics competence is located within the firm – is it embedded in operations (decentralised) or a separate department (centralised). As findings show, a decentralised organisation often mean that analytics staff is deeply engaged in operations and specific solutions. Consequently they usually bring a short-term perspective on machine learning solutions. Locating analytics closer to management and strategy, rather than in operations, will provide a more long-term perspective on solutions. Thus, management support is a significant resource in building this capability because analytics staff must be allowed to get insight into the strategic workings of the organisation. It could, for instance, involve the inclusion of analytics staff in the conversations on organisational goals, future ambitions, investment plans etc. This allows analytics staff to consider how machine learning methods and data could be exploited to achieve set goals.

Developing the capability to connect analytics to strategy also requires a change in mind-set when adopting new machine learning projects. In the studied organisations, machine learning projects often started with the technology itself, following a reasoning similar to: “everybody’s doing machine learning – we should do it too! Let’s find data we can use to apply this cool technology.” Starting with the technology nevertheless hampers the business relevance. Unfortunately the hype around machine learning has created an interest for “robotics” rather than for the core technology. It seems important to steer away from thinking about machine learning as mainly chatbots or intelligent personal assistants, because that is not how machine learning will generate a competitive advantage for companies.

Instead it is important to go back to consider machine learning as a method for finding patterns in large quantities of data. Instead of starting with technology, organisations should start with considering what strategic goals they are striving to achieve. Let the strategic goals determine the mean, and use machine learning as a mean if it serves the purpose. Considering the strategic goals first will ensure that machine learning and analytics are connected to fulfilling a purpose that is relevant to the organisation. The table below summarizes the difference in approach.

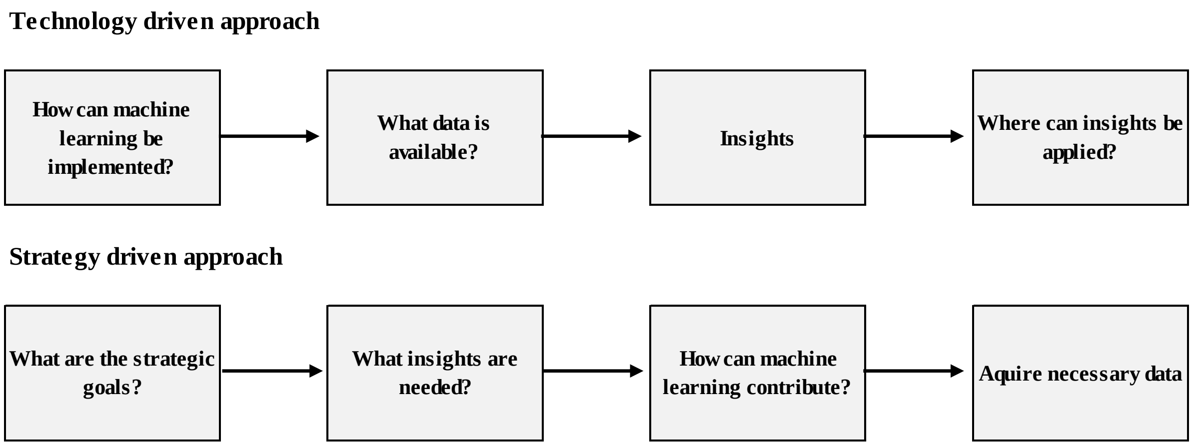

Fig. 12 Technology vs strategy driven approach to adopting machine learning

Educating – Bridging the Knowledge Gaps

The results identified dual inter-organisational knowledge gaps in the studied organisations.
The knowledge gaps refer to the difference in the level of understanding between data scientists (who have deep knowledge about BDA/ML) and management (who have limited or no knowledge), as well as the difference in knowledge been data scientists and the rest of the organisation (who also have limited or no knowledge). Although the analytical framework addressed management’s knowledge gap, the empirical findings support the argument that simply educating management is not enough to drive the adoption and deployment of machine learning. In order to fully exploit the potential of machine learning, there has to be a widespread basic understanding for advanced analytics in organisations.

A first step in developing education capabilities is to acknowledge the knowledge gaps as a result of the inability to educate and inform, rather than a resource gap. Consequently, acquiring talent in machine learning and analytics will not close the gap. The apparent frustration among several interviewees originates from this core issue - the belief that hiring competency in machine learning will result in the adoption and deployment of machine learning. In reality the process is much more complex. Although resources in analytics staff are necessary, closing the gap means that the organisations must develop the capability to spread that knowledge and educate employees in other business areas.

Organisation's Knowledge Gap

A first step towards bridging the organisation's knowledge gap is to educate staff at all levels of the firm. The capability to educate involves a resource combination of analytics staff and business units. It can be debated whether education should be an enriching activity, building on existing resource combinations and capabilities, or if it requires some pioneering activities to develop a new capability. Some of the studied organisations are already trying to educate their organisations, but few are doing it in a highly intentional or systematic manner. As discussed, not having a systematic learning system obstructs the exchange of knowledge and in the long run, innovation. The results show that in the studied organisations most of the education happens ad-hoc by data scientists. However, it doesn't seem reasonable that individual data scientists are, in addition to driving advancements in their designated functions, also responsible for championing the potential of machine learning projects and educating the rest of the organisation. Through enriching activities, such as putting together training guides or an internal training programmes, firms can build on the resources and knowledge they have to develop a more systematic way of communicating and spreading the knowledge on analytics. This sort of initiatives would nevertheless require management's support. To build education capabilities with longevity, management support has to be thrown into the resource mix.

Management's Knowledge gap

Bridging management's knowledge gap would require a resource combination of management support and analytics staff. Education for executives should be adapted to further their understanding of what data scientists and the organisation need from them to manage the adoption and deployment of machine learning. To build the capability to educate management, it seems reasonable to rely on management's support. However, several organisations noted that they are lacking management's support for developing AI and machine learning projects. How do you bridge management's knowledge gap without their support?

The straight-forward answer by the resource-based view would be to look for another resources to combine, and build other capabilities. It can be difficult for one data scientist to go at it alone, and a suggestion would be to create a forum to gather interest around data. This relates back to knowledge creation and the importance of creating forums where tacit knowledge can be exchanged. So going back to the business unit resource. Data scientists and business units are enough resources to build educating capabilities, and they can build capabilities to innovate around machine learning together. It simply comes down to finding a way to integrate people with a common interest to let them exchange knowledge, to create new ideas and understandings. Thus creating 'artificial constructed forums' that bring these different knowledge branches together, which otherwise would not interact, seems key to drive the development of machine learning when organisations cannot rely on management's support. These new capabilities built will hopefully generate momentum around AI and machine learning. If the combination of analytics staff and business units can identify relevant experiments or areas of application for machine learning, empirics support that it will most likely catch the interest, and perhaps also the support, of management.

Breaking-down language barriers

Bridging a gap can be done from two ways. The last section argued the importance for organisations to build capabilities to educate management and the organisation in machine learning and advanced analytics. However, the analytics staff also carry a responsibility in bridging the gap. They have to develop the capability to talk about machine learning and data analytics in non-technical terms. Many terms related to machine learning, analytics and AI are abstract and have different interpretations. Take the term AI for instance, sometimes it is treated as broad term, and sometimes it is considered quite specific.

Relating back to the capability to connect analytics to strategy, it appears that a similar approach to talk about goals rather than technology could facilitate this process. Findings reveal that people often lose interest in analytics when they don't understand the terminology used. Data scientists and analytics staff should thus consider how they communicate around these issues with the rest of the organisation. They can become an instrumental part in bridging the knowledge gaps by adapting the language they use and the approach they have when discussing advanced analytics or machine learning with the laymen in the organisation. Talking about goals, purposes and functions are easier for management and business units to relate to rather than discussing problems in terms of machine learning solutions.

Summary of capabilites to bridge identified gaps		
Gap	**Resource Combination**	**Capability**
Adaptability	Management support + business units + Analytics manager	**Recruiting Influencers:** An effective way of overcoming hold habits and drawback of legacy is to recruit new key talents with the power and influence to traverse old and ineffective habits, routines and cultural norms
Culture	Analytics staff + business units + management support + Analytics manager	**Applying analytics:** Data scientists work with business units and management use let data and analytics drive decisions rather than support them. Apply advanced analytics and integrate insights into strategic decisions.
Effective Organisational Set Up	Analytics staff + business units	**Integrating org. models - the hybrid:** Let analytics staff move between centralised team and the business units, to benefit from the advantages of both the centralised and decentralized model of organising competence.
Data Collection Processes	management support + business units + analytics staff + analytics manager + software	**Governing data**: Improve the collection of data by setting organizations wide standards for how it is collected and managed. Involve business units to ensure everyone knows how to handle data
Talent	management support + analytics manager	**Develop Talent:** It can be difficult to recruit talent with enough business experience, interpersonal skills etc. so develop the capability to develop talent in house through training and integration in business operations
Management's Knowledge	management support + analytics manager	**Educating Systematically** Educate management in the fundamentals of BDA and ML.
Organisation's Knowledge	business units + analytics staff	**Educating Systematically + Brainstorming** Educate and train business units to further their understanding for BDA and ML in a Create the opportunity to data scientists and business units to socialise to enable the communication of tacit knowledge
Business Relevance	analytics staff + management support + analytics manager	**Connecting analytics to strategy:** Integrate the knowledge of management and analytics staff. Include analytics staff in discussions on strategic goals, let analytics drive decisions

Fig. 13 Summary of capabilities to bridge identified gaps

7.4 Additional findings: The Importance of Middle Managers

Drawing from both resource-based theory and innovation theory, it possible to argue for the unique role analytics managers have in the future development of big data analytics and machine learning. Resource-based theory emphasise the mangers' responsibility in combining resources to build new capabilities. However, it requires a deep understanding of the resource's functionality and when it comes to advanced analytics and machine learning, as this study as established, business executives lack this insight. This can explain why so many of the studied organisations were developing machine learning capabilities from the bottom-up. The data scientists at the bottom are the ones with enough insight into the functionality of the necessary resources to be able to see how they can join together to build new capabilities.

However, as this study also reveals, many organisation are appointing analytics managers/department heads with a high skillset in data science and analytics. These individuals have a unique position within the organisation to build the necessary capabilities to adopt machine learning. Not only do they understand the functionality of resources needed in terms of technology, but they also have a unique insight into effective usage areas through their connection with top management. Thus, these individuals will have a unique role in addressing the gaps in vision, knowledge creation and knowledge integration/application. Because of their unique skillset and positon within the organisation they can advise top managers on how to embed big data analytics and machine learning in their long-term strategies and future visions. They are also able leaders in the creation and integration of knowledge exchange as they speak the 'technical language' of the data scientists and the 'business language' of top management. Some organisations have attempted to address this issue by implementing a Chief Digital Officer, with mixing results. Someone in the c-suite will almost always be too distant from the knowledge exchange that occurs at the bottom of the organisation. It leads to the conclusion that the imperative individuals for the adoption of advanced analytics and machine learning are the organisation's middle managers in analytics. It is the growth of a middle-top-down management, where the analytics department managers are the ones who possess the power to influence the future implementation of big data analytics and machine learning.

7.5 Revisiting the analytical framework

The empirical study identified gaps in organisational and managerial attributes, as well as necessary resources and capabilities to bridge those gap. The original analytical framework was constructed based on previous research. With the results of this study it is possible to refine the framework further. The refined framework specifies the necessary resources and capabilities organisations must develop to adopt machine learning in more detail. However, there is room for even more granularity and the framework can be refined further by future studies. The figures below compare the old framework with the new one.

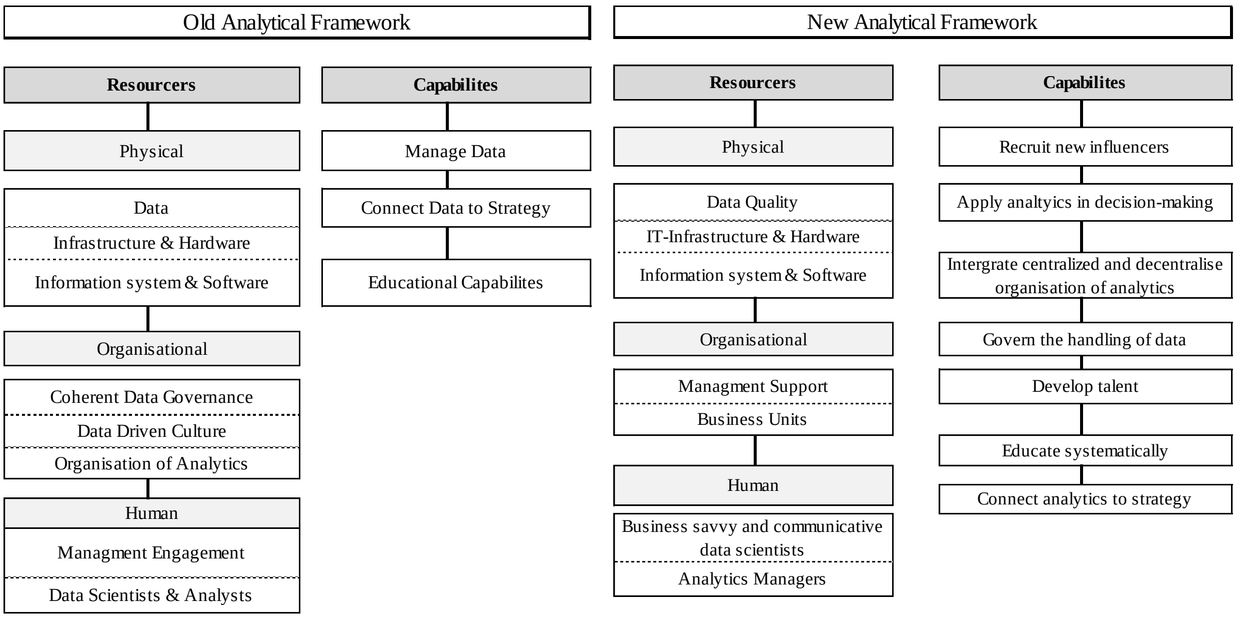

Fig. 14 Comparison of frameworks

Chapter 8. Conclusion

8.1 Theoretical Contributions

By researching and answering the questions (1) *What are the gaps in vision, knowledge creation and knowledge integration/application delaying the adoption of machine learning*? and (2) *What resources and capabilities should organisations develop to bridge the identified gaps?* this study has made contributions to theory and narrowed the research gap on the implementation of machine learning from an organisational and managerial perspective. Through a multi-case research design this study contributes with in-depth insights into the resources and capabilities organisations must acquire and develop to successfully adopt and deploy machine learning. It responds to the demand for more empirical research on how organisations need to change to embrace big data and big data analytics (Gupta and George, 2016; 2015; McAfee et al, 2012).

Furthermore, this study offers a unique perspective on the implementation of machine learning by researching the views of data scientist and analyst at the forefront of these developments. The findings show that these professionals are often the driving force behind the organisations' machine learning initiatives. Thus, they are able to offer first-hand insights into the gaps in resources and capabilities that are hampering advancements. By taking a resource-based-view, this study offers tangible insights on improvement areas to enable the adoption of machine learning. Therefore it contributes to the urgent demand among practitioners for tangible advice on how to enable the implementation of big data analytics (Mikalef et al., 2017).

8.2 Managerial Implications

The findings of this study is of relevance to organisations who have already adopted machine learning for analytics as well as for those who are planning on doing it in the future. The presented insights serve as a foundation for business executives and managers to further investigate and specify gaps in their unique resources and capabilities.

On a general level it can be concluded that organisations that are looking to work with big data, advanced analytics and machine learning should invest in their IT infrastructure, software and hardware. Without the fundamental architecture in place, recording, storing and processing big data will be difficult. It is also important to sort out data bases and investigate the quality and quantity of data available. Data processed with machine learning must align with the problem that is being addressed, thus knowing what data you have and what it means is important.

Although having the data and infrastructure in place for big data analytics is essential, it is of little use if people in the organisation lack the capability to use it. Investing in data scientists and machine learning talent is important, but so is informing the rest of the organisation about the possibilities advanced analytics and machine learning offer. Thus, a significant task for management is to design a plan for the incremental steps the organisation must take to gradually increase the knowledge, talent, skills and interest in big data analytics and machine learning. Naturally, executive management might not have enough insight into the implementation process, and thus the plan should be designed in collaboration with people who do, like data scientists and department heads of analytics. Findings demonstrate how many initiatives in machine learning are passion projects, adopted by individuals because of their interest in the subject. Management also has a responsibility to acknowledge such projects and grow that talent and interest.

A plan for how to increase the analytics maturity of the organisation should also address how analytics can interact with the business areas. Both centralised and decentralised models have benefits and drawbacks, thus it is recommended that organisations elaborate with the idea of applying a hybrid model. Is it possible to let data scientists move between a centralised team and business units, to achieve both "a competence centre" and an integration between analytics knowledge and business insights?

The above recommendations assume the collaboration and support of executive management. As findings demonstrate, not all organisations have managements that are driven, interested or focused on implementing big data analytics and machine learning. In those cases, change must start further down in the organisation and work itself up. In those situations it can be useful to design tools that bring together individuals across the organisation who have an interest in the subject. A forum that allows employees, who otherwise would not have a chance to interact, to discuss and exchange ideas and knowledge on how to adopt and deploy machine learning could be a starting point. If enough momentum can be created around big data, and experiments could be deployed in the organisation that demonstrate how machine learning can drive business value, it is likely to catch the management's interest and win their support.

8.3 Limitations

The generalisation of the findings of this study is limited by its scope and the accessibility to case study organisations. In most organisations only one individual was interviewed, which generates a highly subjective view of the situation. Several interviews should be done to provide a more nuanced perspective of the situation in each organisation. The interviewees also describe the situation from their perspective as professionals working with data science and analytics. Although industry experts (consultants) were consulted for an outsiders' perspective, a more complete picture of the adoption of machine learning in each case could be given my interviewing individuals in top management as well.

The study also attempts to cover a range of industries, to explore what aspects of the phenomenon are universal. There is a limited amount of industries represented, thus the applicability of the results in industries that are significantly different to the ones studies is restricted. The focus of the study has also been on established organisation of a certain size. The findings may not be applicable to start-ups and smaller organisations.

8.4 Future Research

There is a high degree of novelty around the implementation of machine learning and thus future research can take numerous directions. The author sees four interesting paths to build on the findings of this study: (1) Extend the study to incorporate a greater variety of industries, firms and individuals. Extending the study would allow a more complete study of the resources and capabilities needed and identify more refined differences between industries.

(2) The resource-based view offers a static perspective of a firm's resources and capabilities. To understand what resources and capabilities are needed as the business environment is changing over time, it would be interesting to study the phenomenon from a dynamic-capability point of view

(3) Although this study identifies gaps in resources and capabilities that should be addressed to implement machine learning, there are limited results on how to do it. Practitioners demand more tangible suggestion on how to narrow identified gaps. Dwelling further into the organisational resources and capabilities to research efficient ways of narrowing gaps would be an interesting direction of future research.

(4) A final option would be to dwell further into one of the identified gaps and research potential solutions. For instance, the knowledge gap is described as a fundamental challenge for organisations. What should be best practice for increasing the knowledge and interest in big data analytics and machine learning in organisations? Evaluate different approaches to bridge the gap and determine what actions are yielding results and which are not.

List of Works Cited

Armstrong, H. (2015), 'Machines that Learn in the World.' Nesta UK.
< https://www.nesta.org.uk/sites/default/files/machines_that_learn_in_the_wild.pdf>

Barney, J. (1991). 'Firm Resources and Sustained. Competitive Advantage.' *Journal of Management*. 17(1), 99-120

Berner, M., E. Graupner., and A., Maedche. (2014)The Information Panopticon in the Big Data Era. Journal of Organization Design, 3(1), 14-19

Bharadwaj, A.S. (2000). 'A resource-based perspective on information technology capability and firm performance: An empirical investigation.' *MIS Quarterly*, 24 (1), 169-196

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., and Venkatraman, V. (2013). 'Visions and Voices on Emerging Challenges in Digital Business Strategy,' *MIS Quarterly*. 37 (2), 633-661

Bedi et al. (2017). 'The Global CIO point of View.' Servicenow.
<https://www.servicenow.com/content/dam/servicenow/documents/whitepapers/wp-cio-global-pov.pdf>

Bughin, J., E. Hazan, S. Ramaswamy. M. Chui., Tera Allas., P Dahlström. (2017). 'Artificial Intelligence, The Next Digital Frontier.' *McKinsey Global Institute Study*.

Brinkhues, R, Maçada AC, Casalinho G. (2014). Information management capabilities: antecedents and consequences. In: Proceedings of Americas conference on information systems (AMCIS)

Bryman, A., & Bell, E. (2015). Business research methods. Oxford University Press, USA.

Brynjolfsson, E, & L. Hitt., (2000). 'Beyond Computation: Information Technology, Oragnizational Transofmration and Business Performance.' Journal of Economic Perspectives, 4 (14), 23-48.

Brynjolfsson, E, E. Renshaw and M. Van Alstyne,. (1996). 'The Matrix of Change: A tTool for Business Porcess Reengineering. Working Paper Series,' *MIT Center for Coordination Science*.

Cambridge Analytica. (2017). 'Cambridge Analytica Case Studies.' *Cambridge Analytica*.
<https://ca-political.com/casestudies>

Chen, Hsinchun; Chiang, Roger H. L. and Storey, Veda C. (2012). 'Business Intelligence and Analytics: From Big Data to Big Impact.' *MIS Quarterly*, 36 (4), 1165-1188

Davenport, T. H., Barth, P., & Bean, R. (2012). 'How 'big data' is different.' MIT *Sloan Management Review*, 54, 43–46

Davenport, T. (2006). 'Competing on Analytics,'*Harvard Business Review*. Jan, 99–107.

Davenport T. & D.J Patil. (2012). 'Data Scientist the Sexiest Job of the 21st Century'. *Harvard Business Review*. Oct., 6- 12
< https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

Denscombe, D. (1998). The good research guide for small-scale social research projects, Buckingham: Open University Press.

Dhar, V. (2013). ‘Data science and prediction.’ *Communication of the ACM.* 56 (12), 64-73.

Dykes, B. (2017). ‘Creating a data-driven culture: Why leading by example is essential,’ *Forbes*. < https://www.forbes.com/sites/brentdykes/2017/10/26/creating-a-data-driven-culture-why-leading-by-example-is-essential/#631f7a336737>

Dong, C. (2017). ‘The Evolution of Machine Learning.’ TheCrunchNetwork. <https://techcrunch.com/2017/08/08/the-evolution-of-machine-learning/>

Dubois, A. & L.K Gadde. (2013). ‘Systematic-Combining – a decade later.’ Journal of Business Research, 67 (6), pp. 1277-1284

Drucker, P.F. (1993). *Post-capitalist Society*. New York: Butterworth Heineman

Erlandson, et al. (1993). *Doing naturalistic inquiry: a guide to methods,* London: Sage

Espinosa J, & F, Armour. (2016) "The Big Data Analytics Gold Rush: A Research Framework for Coordination and Governance", 2016 49th Hawaii International Conference on System Sciences (HICSS), 1112-1121

Flick, Y. (2014). An Introduction to Qualitative Research (5th ed.). London: Sage.

Fransman, M. (1990). *The Market and beyond, co-operation and competition in information technology in the Japanese system*. Cambridge: Cambridge University Press.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. & Gnanzou, D. (2015). How 'big data' can make big impact: findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165 234-24

Galbraith, Jay R. (2014). ‘Organizational Design Challenges Resulting from Big Data’. *Journal of Organization Design*, 3(1), pp. 2-13

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

Garmaki M, Boughzala I, Wamba SF. (2016). ‘The effect of big data analytics capability on firm performance.’ In: Proceedings of 20th Pacific Asia conference on information systems, PACIS

Germann, Frank, Gary L. Lilien Lars Fiedler and Matthias Kraus (2014). ‘Do Retailers Benefit from Deploying Customer Analytics?’ Journal of Retailing. 90 (4), 587-593

Gibbs, S. (2014). ‘Google buys UK start-up Deepmind for £400m.’ *The Guardian*. <https://www.theguardian.com/technology/2014/jan/27/google-acquires-uk-artificial-intelligence-startup-deepmind>

Grant, R. M., (1991) ‘The resource-based theory of competitive advantage: implications for strategy formulation,’ California Management Review 33 (3), 114-135

Gregshorn, D. (2017). ‘The industry that predicts your vote – and alters it – is still just in its infancy.’ *Quartz Media*. <https://qz.com/977429/the-industry-that-predicts-your-vote-and-then-alters-it-is-still-just-in-its-infancy/>

Grossman, R. & K, Seigel. (2014). ‘Organisational Models for Big Data Analytics.’ *Journal of Organisational Design*. 3 (1), 20-25

Größler, A & A, Grübner. (2006). ‘An empirical model of the relationships between manufacturing capabilities,’ *International Journal of Operations & Production Management*, 26 (5), 458-485

Gomez-Uribe C. & Neil Hunt. (2016). ‘The Netflix Recommender System: Algorithms, Business Value, and Innovation.’ *ACM Trans. Manage. Inf. Syst.* 6(4) article 13.

Gupta M, & JF, George. (2016). Toward the development of a big data analytics capability.’ *Information Management* 53(8), 1049–1064

Hatch, M. (2012). *The Maker Revolution: Building a future on creativity and innovation in an exponential world*. John Willey & Sons. USA.

Helfat, C.E. & Lieberman, M. (2002). ‘The birth of capabilities: market entry and the importance of pre-history.’ *Industrial and Corporate Change*, 11(4), 725-760

International Institute for Analytics. (2016). *IIA Business Intelligence and Analytics Capabilities Report*, IIA & SAS Institute.

Jacobi F, Jahn S, Krawatzeck R, Dinter B, Lorenz A. (2014). ‘Towards a design model for interdisciplinary information systems curriculum development, as exemplified by big data analytics education.’ In: Proceedings of European conference on information systems (ECIS).

Johannessen, J. A., Olsen, B., & Olaisen, J. (1999). ‘Aspects of innovation theory based knowledge management.’ *Journal of International Management*, 19(2), 121–139

Jones, T.M. (2017). ‘A beginners guide to Artificial Intelligence, machine learning and cognitive computing.’ *IBM Developer Works*.
< https://www.ibm.com/developerworks/library/cc-beginner-guide-machine-learning-ai-cognitive/index.html>

Kamioka, T. and T, Tapanainen. (2014). ‘Organizational use of big data and competitive advantage—exploration of antecedents.’ *In: Pacific Asia conference on information systems* (PACIS)

Lane, A. (2017). “We are likely 3-5 years out from advanced analytics being table stakes and critical to the viability of a company.” The Innovation Enterprise.
<https://channels.theinnovationenterprise.com/articles/we-are-likely-3-5-years-out-from-advanced-analytics-being-table-stakes-and-critical-to-the-viability-of-a-company>

Lamba, HS. & SK Dubey. (2015). 'Analysis of requirements for big data adoption to maximize IT business value in reliability.' In: 2015 4th International conference on infocom technologies and optimization (ICRITO) (trends and future directions). IEEE, 1–6

Lanely. D (2001). 'Application Delivery Strategies.' *MetaGroup*. < https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

LaValle, S, E. Lesser, R. Schockely, M.S. Hopkins and N. Kurschwitz. (2011). 'Big Data, Analytics and the Path from Insight to Value,' MIT Sloan Management Review. 52 (2).

Lee, R.P & R. Grewal. (2004). 'Strategic Responses to New Technologies and Their Impact on Firm Performance.' *Journal of Marketing*. 68(4), 157-171

Leech, B. L. (2002): Asking Questions: Techniques for Semi structured Interviews. In: Political Science and Politics (35) 4, 665-668

Leonard-Barton, D.L. (1995). *Wellsprings of knowledge: Building and sustaining the sources of innovation*. Boston, MA: Harvard Business School Press.

Lincoln, Y.S & E.G. Guba, (1985). Naturalistic inquiry, Beverly Hills: Sage.

Liu, Y. (2014). Big data and predictive business analytics. The Journal of Business Forecasting, 33, 40–42.

Lockett, A., Thompson, S. and Morgenstern, U., (2009). 'The development of the resource-based view of the firm:a critical appraisal.' *International Journal of Management Reviews*, 11 (1), 9-28.

Manyika J, Chui M, Brown B, Bughin J, Dobbs B, Roxburgh C, Byers AH. (2011). 'Big data: The next frontier for innovation, competition, and productivity'. *McKinsey Global Institute*. <htp://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation>

Mata, F.J., Fuerst, W.L. and Barney, J.B. (1995). 'Information Technology and Sustained Competitive Advantage: A. Resource-Based Analysis.' *MIS Quarterly*, 19, 487-505

McAfee A, Brynjolfsson E. 2012. Big data: The management revolution. *Harvard Business Review* October 2012: 1–9.

McCarthy J, Minsky ML, Rochester N, and Shannon CE. (1955). 'Proposal the Darthmouth summer research project on artificial intelligence.' Technical report, Dartmouth College, Hanover (NH).

Melville, N., Kraemer, K. and Gurbaxani, V. (2004). 'Information Technology and Organizational Performance: An Integrative Model of IT Business Value.' *MIS Quarterly*, 28, 283-322

Merriam, S.B (1998). Qualitative research and case study applications in education, San Francisco: Jossey-Bass

Netflix Media Centre. (2017). "About Netflix". *Netflix* <https://media.netflix.com/en/about-netflix>

Mikalef, P., Pappas, I. O., Krogstie, J. and Giannakos, M. (2017). 'Big data analytics capabilities: a systematic literature review and research agenda.' *Information Systems and e-Business Management*.

MIT Sloan. (2016). 'Lessons From Becoming a Data-driven Organisation.' *MIT Sloan Management Review*.

Murnane R. J, F. Levy, D.H. Autor. (1999). 'Technological Change, Comptuers and Skill Demands: Evidence from the Back Office of a Large Bank.' Harvard University Graduate School of Education.

Parloff, R. (2016). 'Why deep learning is suddenly changing your life.' *Fortune*. < http://fortune.com/ai-artificial-intelligence-deep-machine-learning/>

Pearsall. J (Ed.) (2001). *The New Oxford Dictionary of English*, Oxford University Press,

Penrose E.T. (1959). The Theory of the Growth of the Firm. Oxford: Blackwell

Peteraf, M. (1993). 'The cornerstones of competitive advantage: A resource-based view Strategic Management Journal', 14(3),179-191.

Peteraf, M.A. & Bergen M.E. (2003). 'Scanning Dynamic Competitive Landscapes A Market-Based and Resource-Based Framework.' *Strategic Management Journal*, 24, 1027-1041

Provost F. and T. Fawcett. (2013) Data Science and its Relationship to Big Data and Data-Driven Decision Making. Big Data, 1(1), 51-59.

SAS. (2017). Machine Learning: *What it is and why it matters*. < https://www.sas.com/en_ie/insights/analytics/machine-learning.html>

Rosengren, Lina. (2017). 'Stena Lina storsatsar på AI, "en fråga om överlevnad"', *CIO*. <https://cio.idg.se/2.1782/1.681823/stena-line-ai-fraga-om-overlevnad>

Ren S, Wamba SF, Akter S, Dubey R, Childe SJ. (2016). 'Modelling quality dynamics, business value and firm performance in a big data analytics environment.' *Int. Journal of Prod. Res*. 55(17), 5011–5026

Rubin, H. J., & Rubin, I. S. (2011). *Qualitative interviewing: The art of hearing data*. Sage.

Schroeck, M., R. Shockley, J. Smart, D. Romero-Morales, P. Tufano. (2012). 'Analytics: The real-world use of big data. How innovative enterprises extract value from uncertain data.' IBM Institute for Business Value.

Shalev-Shwartz. S. & Shai Ben-David. (2014). *From Theory to Algorithms*. Cambridge University Press. UK.

Sharma, R., Mithas, S. and A. Kankanhalli, (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, 23 (4), 433-441

Shenton, A.K. (2004). ‘Strategies for ensuring trustworthiness in qualitative research projects.’ *Education for Information.* 22, 63–75

Silverman, D. (2016). *Qualitative Research*. (4th ed.) London: Sage.

Siri Team. (2014). ‘Deep Learning for Siri’s Voice: On-device Deep Mixture Density Networks for Hybrid Unit Selection Synthesis.’ *Apple Machine Learning Journal* (1) 4. (online) < https://machinelearning.apple.com/2017/08/06/siri-voices.html>

Spencer L, Ritchie J, Lewis J, Dillon L .(2003). *Quality in Qualitative Evaluation: A framework for assessing research evidence*. London: Cabinet Office

Stewart, T.A. (1997). Intellectual capital: [1]he new wealth of organizations. London: Doubleday.

Sirmon, D.G., Hitt, M.A., & Ireland, R.D. (2007). ‘Managing firm resources in dynamic environments to create value: Looking inside the black box.’ *Academy of management review*, 32(1), 273-292

Teece, D. & G. Pisano. (1994). ‘Dynamic Capabilites of the Firms – an Introduction.’ *Industrial and Corporate Change*. 3 (3), 537–556
Teece, D. Pisano G., and Shuen A. (1997). *Dynamic Capabilities and Strategic Management*. Strategic Management Journal, 18, 7, 509-53

Tusman, M., & Nadler, D. (1986). ‘Organizing for innovation.’ *California Management Review*, 28 (3), 74- 92.

Tweney, D. (2013). ‘Walmart scoops up Inkiru to bolster its ‘big data’ capabilities online’ [online].

Ulaga, W. & R, Werner. (2011). ‘Hybrid Offerings: How Manufacturing. Firms Combine Goods and Services Successfully,’ *Journal of Marketing,* 75 (6), 5-23

Vidgen R, Shaw S, Grant DB. (2017). ‘Management challenges in creating value from business analytics.’ *European Journal on Operational Resources*. 261(2), 626–639

Wedel. M, & P.K. Kannan. (2016). ‘Marketing Analytics for Data-Rich Environments,’ *Journal of Marketing,* 80 (6), 97-121

Wernerfelt, B.(1984). ‘A Resource-based View of the Firm,’ *Strategic Management Journal*, 5 (2), 171-180

Winters, D. (2015). “What is the difference between data analytics, data analysis, data mining, data science, machine learning and big data?”. Quora. <https://www.quora.com/What-is-the-difference-between-Data-Analytics-Data-Analysis-Data-Mining-Data-Science-Machine-Learning-and-Big-Data-1>

Conference: Analyticsdagarna 2017 - 11/10/2017 in Stockholm

Appendix

Appendix 1: Big Data – The other four “Vs”

In addition to Volume, Velocity and Variety, some scholars have added other distinctions to the conceptualization of big data. *Veracity* is the fourth “V” commonly acknowledged to definite big data. It refers to the uncertainty of data and the degree to which it is protected from biases, noise and modification (Demchenko et al., 2013). Ensuring the quality and reliability of big data is essential if it is being analysed and applied to drive business value (Akter et al., 2016). Consequently, some researchers suggest that data should be authenticated and pass quality-compliance procedures before being applied in business decisions (Dong & Srivastava 2013; Ganomi & Haide, 2015).

More recent literature on big data extends the four “Vs” to include: value, variability and visualization. *Value* refers to the extent big data generate insights that drive economic benefits for a firm (Fosso Wamba et al., 2015). A challenge with big data is its low value density, meaning that the value from processing large data sets is proportionally low compared to its volume (Mikalef et al., 2017). *Variability* refers to how the meaning and insights of data are constantly changing as the same information is interpreted in different ways (Seddon and Currie, 2017). It is particularly evident in language processing, as words do not have static definitions and their meaning can change depending on context. For instance, a programme interpreting reviews must be sophisticated enough to understand that a word such as “great” can refer to both positive and negative experiences (“I had a great experience at your restaurant” versus. “I am greatly disappointed at your service”). Finally, *visualization* describes the process of interpreting patterns and trends in the data with the help of artificial intelligence methods. The findings must furthermore be presented, or visualized, in a manner that is readable and accessible (Seddon and Currie, 2017).

Appendix 2: Interviewee Guide

Guide of interviewed individual and when interview took place. Participants we granted anonymity.

Interviewee Guide

#	Position	Industry	Point in Time
1	Director of Sales	Telecom	Sept. 2017
2	IT-Consultant/Data Scientist	IT Consulting	Sept. 2017
3	Solution Manager	Technology	Sept. 2017
4	Head of Digital Innovation Centre	Recruitment	Oct. 2017
5	Data Scientist	Recruitment	Oct. 2017
6	Data Scientist	Energy	Nov. 2017
7	Head of IT Department	Energy	Nov. 2017
8	Data Scientist	Media	Oct. 2017
9	Head of Group Data, BI & Analytics	Media	Oct. 2017
10	Head of Customer Development	Gambling	Oct. 2017
11	Head of Analytics	Gambling	Oct. 2017
12	Data Scientist	Gambling	Nov. 2017
13	Head of Analytics	Gaming	Oct. 2017
14	Data Scientist	Gaming	Nov. 2017
15	Advanced Analytics Manager	Automotive	Nov. 2017
16	Head of Development BI & Analytics	FMCG	Oct. 2017
17	Analytics Manager	Travel	Nov. 2017
18	M.D & Strategic Development Manager	Health Care	Dec. 2017
19	Senior Analytics Consultant	Consultant	Dec. 2017
20	Marketing Manager	FMCG	Oct. 2017
21	CEO	Retail	Nov. 2017

Fig. 15 Interviewee guide

Appendix 3. Identified Gaps in Participating Organisations

An overview of the most noteworthy gaps in each of the organisations.

Industry	Identified gaps in:
Telecom	**Adaptability** - long legacy and inflexible structures **Data Collection** - no unified system, separate silos
Gambling	**Management Support** - management uninterested in AI/ML projects **Effective Org. of Competence** - a centralised model, data scientists don't interact with business operations **Business Relevance** - innovate around AI/ML but project lack business relevance
Gaming	**Talent** - challenge to keep and develop talent **Data collection** - doesn't have storage for mobile data, must increase capacity and improve infrastructure
Gaming	**Talent** - challenge to keep and develop talent **Connecting Strategy to Data** - data support decisions rather than drive them
Media	**Adaptability** - strong sales culture, long legacy, no interest in AI/ML **Management Support** - management uninterested in AI/ML projects
Media	**Organisation's Knowledge** - organisation doesn't know how to integrate analytics or ML **Management Support** - management uninterested in AI/ML projects
Energy	**Connecting Strategy to Data** - data support decisions rather than drive them **Effective Org. of Competence** - a centralised model, data scientists don't interact with business operations **Data Collection** - no unified system, separate silos
FMCG	**Management's Knowledge** - challenge to keep and develop talent **Data Collection** - treatment of data in the past results a lot of time spent on "cleaning data" due to poor data quality
Travel	**Data Collection** - spent a lot time cleaning data, due to using old data sets **Connecting Strategy to Dat**a - data support decisions rather than drive them
Health Care	**Organisation's & Management's Knowledge** - Doctors and Nurses are reluctant to incorporate new technology, few are interested in learning more about new technologies, hospital management are unfamiliar with new technology and cannot initiate projects
Automotive	**Data Collection** - different types of unreliable data sources, spend a lot of time cleaning data **Talent** - challenge to keep talent, experienced staff is highly sought after, make it difficult to retain them
Retail	**Organisation's Knowledge** - Management is driving in becoming more data driven, organisation lack experience - take time to change mind-set to integrate data in decision-making processes
Technology	**Adaptability** - long legacy, a lot of prestige, set routines and frameworks make it hard to innovate around ML **Connecting Strategy to Dat**a - management does not want to be challenged by data
Consultant	**Adaptability** - firms legacy get in the way of innovation **Organisation & Management Knowledge** - large discrepancies in organisation which cause tensions
Consultant	**Adaptability** - firms legacy get in the way of innovation **Organisation & Management Knowledge** - large discrepancies in organisation which cause tensions
Recruitment	**Management's Knowledge** - management is uninformed about the potential of new technology **Talent** - don't have enough talent in data science and ML to implement large scale projects (only experiment)

Fig. 16 Top identified gaps in each industry

Appendix 4. Interview Guide

This is an example of the interview guide. It provided an outline for the interviews but naturally divergences were made from this.

Introduction

- Introduce myself and the purpose of the project
- Ask if they are okay with recording, interview will last approx. 1 hour, confirm anonymity and confidentiality

Background:

- Could you tell me about your roll as X here at company Y?
- What is your background?
- What are the greatest challenges of your job today?

General discussion about AI and machine learning in the industry
First: Explain what I mean by AI – do they agree with this definition?

- How would you describe the advancements in AI in your industry at the moment?
- What challenges are you seeing?
- What are the opportunities?

Discussion focused on the organisation

- Tell me about the machine learning/AI project you are working on?
- How did it come about?
- Who is involved?
- What are the challenges?
- How is the project expected to drive business value?
- What are the expectations?

Discussion on data and analytics

- How are you working with data and analytics in decision-making today?
- What are the differences in how data and analytics is applied within the organisation?
- What are the challenges?

Discussion about the future

- Where do you see your machine learning project in the next 1-2 years?
- Where do you see your organisation relating to BDA and ML in the future?
- What is your role in getting the organisation there?

Company Specific question

Here I would generally make notes about something I had read specifically about the company, the project they were working on or their industry. I would write the questions/topic down to integrate into the conversation when it was suitable.

Appendix 5. Advanced analytics

The following model describes the different forms of analytical methods with increasing complexity from left to right (Wedel & Kannan, 2016). Advanced analytics generally refers to predictive and prescriptive analytics, involving prediction models, machine learning and optimization models.

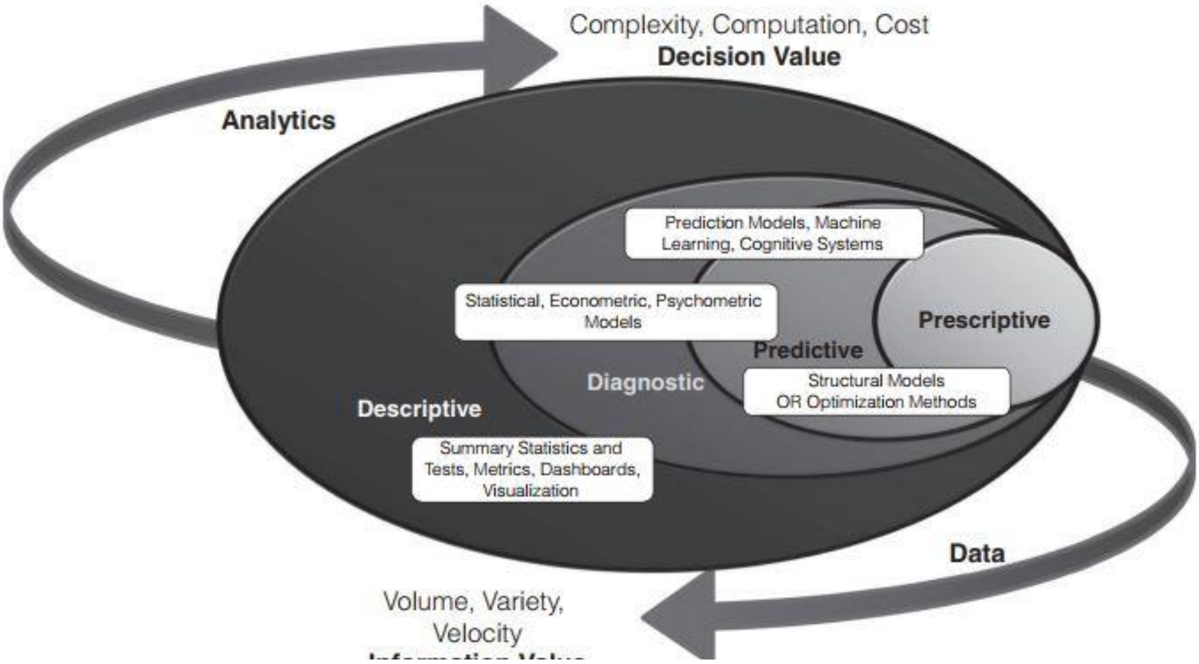

Fig. 17 Different types of analytical methods

The diagram below illustrates the progression of analytics from descriptive to prescriptive. Advanced analytics generally refers to the last two step of this ladder (Jones, 2017).

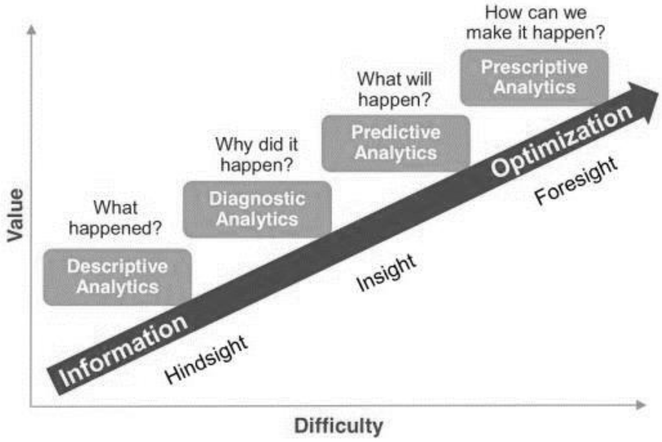

Fig. 18 The analytics ladder

Appendix 6: Illustration of Data Management

The flowchart below illustrates the working relationship between data management and data analysis, as well as the role of visualisation. It clearly demonstrates the importance of good data management in order to produce useful analytics. It further illustrates the key role data visualisation has in bridging data management and data analysis.

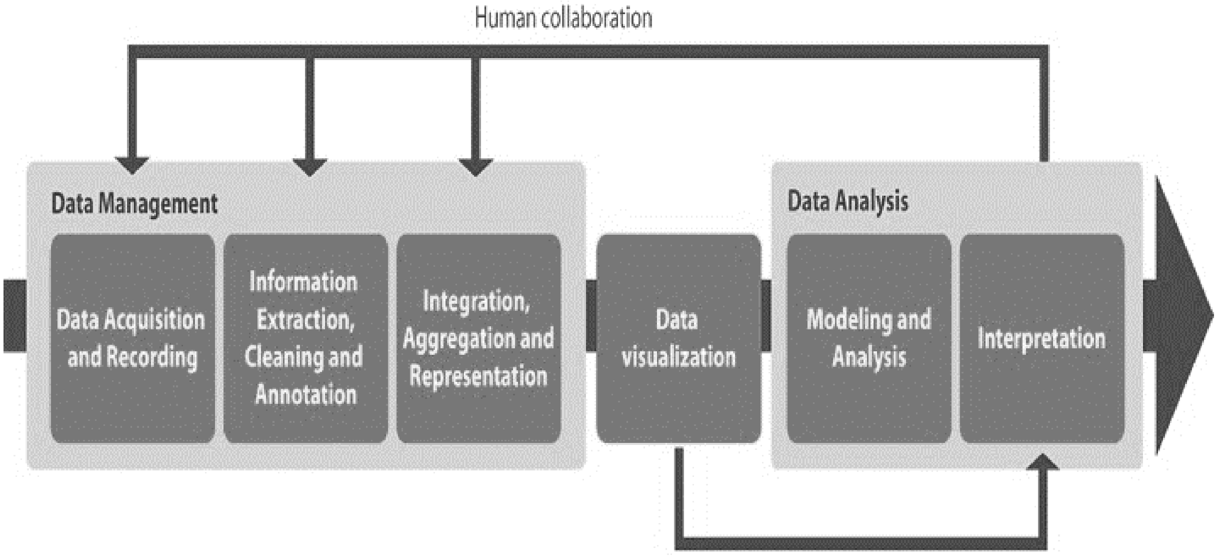

Fig. 19 Data Management Process

Appendix 7: The Machine Learning Process

The figures illustrate the machine learning process, describing the process of from gathering training data, to building the model and generate insights. The process starts with training data and the building of a model. The model are then fed input data to make predictions. The predicted outcomes are evaluated and the model refined accordingly.

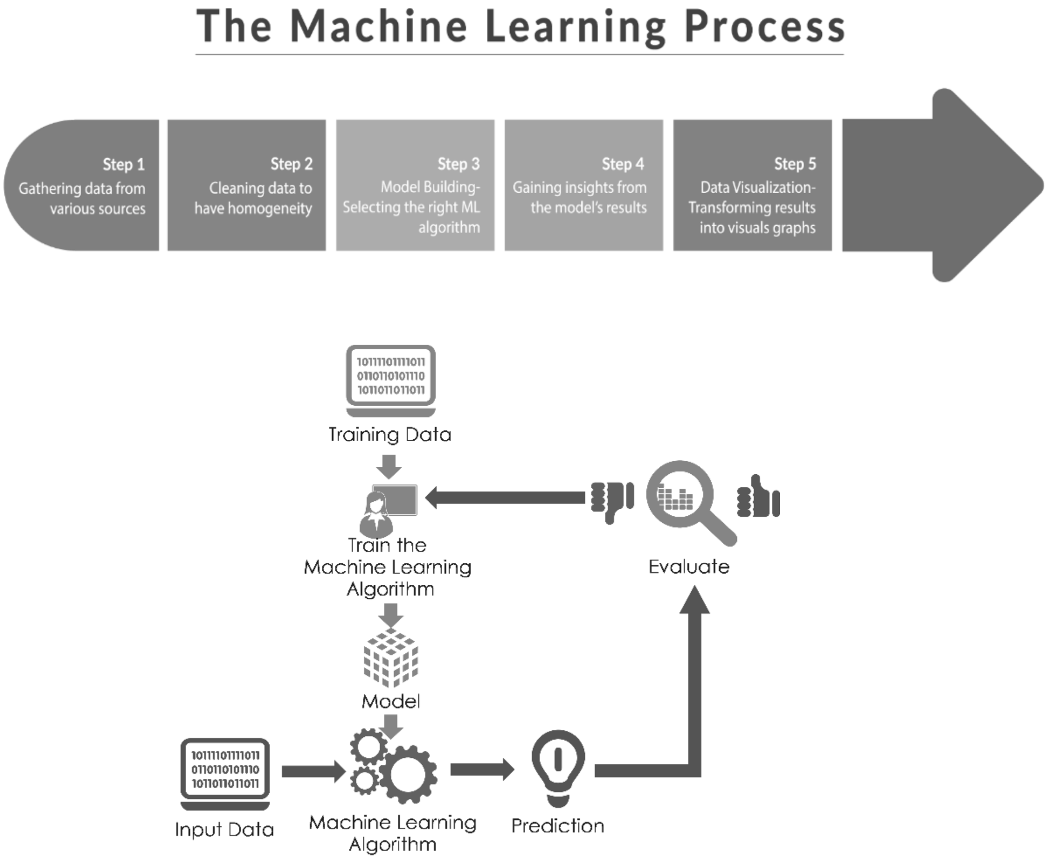

Fig. 20 The machine learning process

Appendix 8: Different Machine Learning Techniques

There are numerous methods to machine learning, and although the specifics is beyond the scope of this thesis, a short overview of the techniques will facilitate the understanding of the widespread areas of application. Machine learning techniques can be categorized into two types of learning, supervised learning and unsupervised learning. In *supervised learning* an AI is given input and the expected output. If there is an error in the output generated, the AI will adjust its calculation and repeat the process until it generates the expected output without errors. An example is predicting the weather with AI. The AI is fed historical data on pressure, humidity, wind etc. and is expected to predict the temperature.

Unsupervised learning, on the other hand, means that the AI is using data without a set framework from the programmer. The AI is expected to group and interpret the available data. An example is predicting the behaviour of online consumers. There are no given inputs and outputs for the AI to use in order to detect which consumers are buying what products. Instead the AI will create its own groups based on the available data and predict which customers are most likely to buy what products.

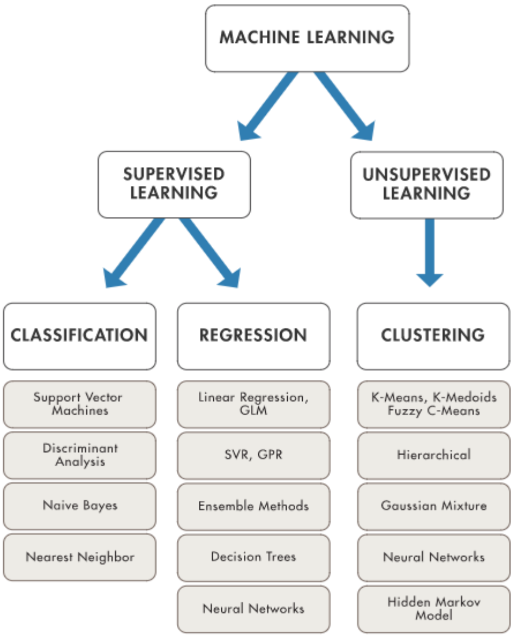

Fig. 21 Different types of machine learning techniques

Appendix 9: AI Index

The AI index indicate the adoption rate of AI technologies across different industries.

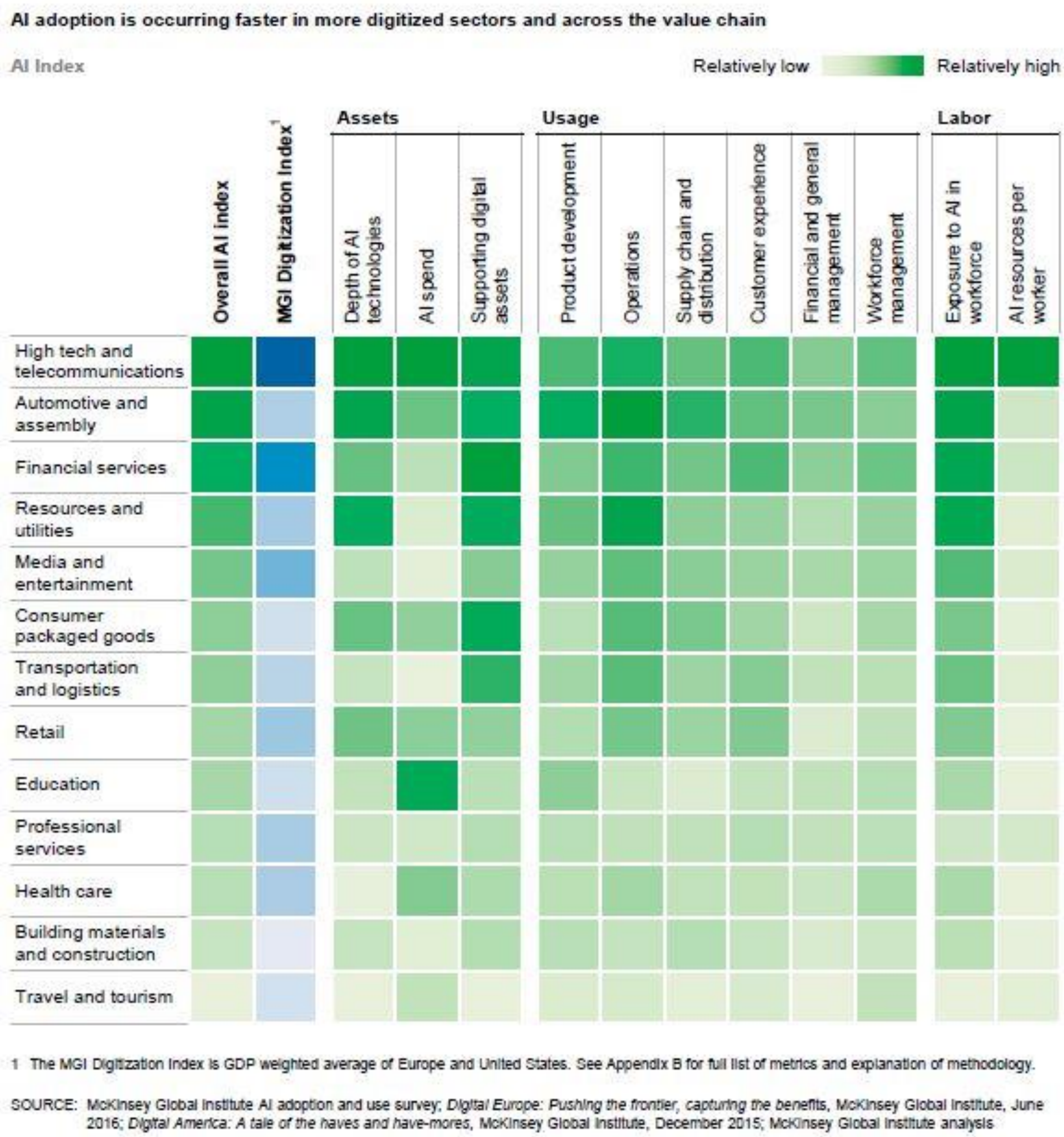

Fig. 22 AI Index