A Field Lab project presented as a part of the requirements for the Award of a master's degree in International Management from Nova School of Business and Economics

NOVA SCHOOL OF BUSINESS & ECONOMICS

Big Data and its Implication on Supply Chain Management

Henrik Svankjær Sommerin, 4103 / 30234

This project was carried out during the Master in International Management Program, with supervision of José Crespo de Carvalho.

4th of January 2019

Acknowledgement

This field lab illustrates the end of my academic journey at Nova School of Business and Economics. I therefore wish to express my gratitude to all my previous professors, with special thanks to José Crespo de Carvalho who has supervised and supported me through this final project. Furthermore, I would like to thank all my interview participants who strongly supported me with their professional insights and experiences on the elaborated topic. Finally, I would like to thank my family, girlfriend, friends and fellow students who made my masters studies an enriching and memorable experience.

Lisbon, 4th of January 2019

Henrik Svankjær Sommerin

Table of Contents

1. Introduction

Internet and other data networks are continuously growing the physical world. This is mainly since physical objects are interconnected via information technology systems (Bauer, 2015). The resulting effect is impacting all areas in life, which will be identified by the work of this thesis. With digital technologies connecting the global economy, data has become a torrent flowing into every area (Economist, 2010). Companies are collecting voluminous information about their entire value chain, including, but not limited to, their customers, suppliers and operations. This data collection is supported by sensors collecting data through physical devices such as mobile phones, wearables, vehicles, home appliances and industrial machines. Simultaneously, data is collected through non-physical processes in the web. Meaning, consumer behaviour can be tracked and identified through purchasing, chatting, sharing and searching patterns. Thus, the physical and non-physical data collection is continuously expanding into new areas, providing data with greater variety, increasing volumes, arriving at ever-higher velocity (Oracle, u.d.). As a result, data sets are too large or complex for traditional data-processing application software to adequately deal with. These data sets are referred to as big data.

Collected data can be used both, isolated, and in combination with other datasets, to improve managerial decision-making. Until recently, complex analytic capabilities delivered close to zero value. However, disruptive technologies, including cloud technologies and machine learning, have enabled big data usage, whereas organisations can position themselves to achieve larger strategic and operational goals (Thomas H. Davenport, 2013).

Furthermore, a greater amount of information availability can very well be aligned with risk and return. First and foremost, there are technological and practical questions to be examined, but sensitive complicated aspects like privacy, security, data governance and human errors can lead to intricate challenges. Thus, big data legislation has granted a completely new dimension for moral and ethics (King & Richards, 2014). Nonetheless, using big data appropriately, can provide organisations a tremendous up-side and increase competitive advantage. In addition, it may arguably be a necessity for long-term competitive advantages.

In terms of generating data, ETL devices (extract, transform, load) are retrieving data and feeds it into databases, either directly or through other databases. Access to internet is a critical aspect in data generation. As an increasing number of interconnected devices are connected to internet, companies can retrieve new and different information. After generating data, it is aggregated in one or more databases, or so-called data marts. This media can exist through multiple resources and be operated on-premise or in the cloud. In the second part of the value chain, tools providing business intelligence, analytics and insights through dashboards, reports, visualising of data and other analytical tools are used (SAS, u.d.).

In addition to size, data can be categorized as structured or unstructured. Structured data is organised into a formatted repository, typically a database, so that its elements can be made addressable for more effective processing and analysis. On the other hand, unstructured is “everything” else, with the likes of pictures, sound files, videos, e-mail, documents and other unorganised data. In short, it does not have a pre-defined data model or is not organised in a pre-defined manner. However, big data and analytics is relying on both data types. To facilitate predictive and prescriptive analytics, both must be consolidated to a coherent one. When done appropriately, these data insights can provide comprehensive descriptions, explanations, predictions and prescriptions on behaviour. On the other hand, decision making done on poor quality may equally become a corporate burden.

2. Research Approach

2.1 Problem definition

This paper takes a practical approach to examine big data and its future implications on the end-to-end supply chain, by investigating the potentialities and limitations within key supply chain levers. Firstly, big data definition, sources, techniques and industry implication will be described to provide a general understanding of the technology. Secondly, status quo will be investigated through secondary data sources to establish a fundamental understanding of big data and analytics. Thirdly, qualitative interviews will be addressed to verify previous findings. The conclusive goal is to understand the implication on current and future supply chain management and to what extend big data and analytics can enhance these activities.

2.2 Methodology

When conducting a scientific paper, two research methods can be applied, inductive- and deductive reasoning. Firstly, inductive reasoning is initialised with detailed observations on the topic, moving towards more abstract generalisations and ideas (Neuman, 2003). The intent is to explain causation based on actual findings. Secondly, a deductive approach is concerned with deducting conclusions from existing theory based on new empirical findings. This approach offers the possibility to explain causal relationships, measure concepts quantitatively and generalise research findings. This paper is based on inductive research, to identify the current and future application of big data analytics as a digital layer in supply chain management. Initially, secondary research was performed to identify the current state of big data analytics, both in general enterprise functions and within supply chain management. Secondly, it was imperative to perform interviews with field experts to validate prior findings. When conducting research on a topic with scarcity of sources, bound to subjectivity and individual reasoning, risk is accepted, as no theory may emerge at all. Consequently, qualitative interviews were highlighted as crucial to identify and validate causation between subjective reasoning and empirical findings. These interviews amounted to a total of three, whereas all were data scientists, respectively working within consulting and banking. The literature review and qualitative research forms the basis of the concluding big data analytics implications on current and future supply chain management. This paper consists of both primary and secondary sources. Primary sources are represented in the qualitative interviews, and have been used to validate big data as a layer in digital supply chains, as well as its pros and cons. The secondary data sources were collected from literature, research papers and other published articles.

2.3 Research questions

Aligned with the abovementioned problem definition this paper aims to investigate three formal research questions.

1. How can big data and analytics be defined and applied?
2. What are the limitations and opportunities?
3. How can big data and analytics be applied within supply chain management levers?

3. General application of Big Data and Analytics

3.1 Literature review

3.1.1 Definition

Big data has been variously defined in literature. However, big data is commonly defined by a suite of three key traits: volume, velocity and variety, constituting the 3V's (Laney, 2001). Building on the 3V's, the most widely adopted definition is that big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making and process automation (Gartner, u.d.).

Compared to a long tradition of capturing transactional data, organisations are nowadays capturing additional data from their operational environment at an increasingly higher speed, such as web data, text data, time and location data, smart grid and sensor data, social network data and so forth.

3.1.2 Data sources

One can suggest three techniques to differentiate big data from traditional data sources (Franks, 2012). Firstly, the author illustrate how big data can be an entirely new source of data. For example, with e-commerce, what separate transaction data from traditional transaction data, is how big data enable the e-commerce provider to capture browsing behaviour to determine purchasing behaviour. Secondly, it is argued how the speed of data feed has increased to an extent where its qualified as a new data source. Also, gradually moving from manually performed tasks, to a more automated society, the frequency at which data is created has enabled a different and more analytic level of data. Thirdly, data is increasingly becoming more semi-structured and unstructured, compared to the structured realm of traditional data. Structured data refers to information with a high degree of organisation, such as accounting information or spreadsheets. Contradictory, unstructured data sources have little or no organisation. This format can be categorized as text-, audio- and video data. One problem with unstructured data, is how it may not be meaningful to the organisation. Thus, the queries executed against unstructured data can produce results that are heavily or negatively influenced by different dialect or languages (Inmon & Krishnan, 2010). Semi-structured data can be defined as irregular or incomplete data sources, with a structure that may rapidly change and is constrained to some degree of unpredictability but does nonetheless contain an underlying logic (Buneman, 1997).

3.1.3 Techniques

Big Data brought with it a so-called paradigm shift, moving from descriptive analytics to an increasing focus on predictive and prescriptive analytics. Descriptive analytics is used to understand past and current business performance and make informed decision. In its essence, it shall allow managers to understand the actual impact of their decision. For example, whether a marketing campaign proved beneficial, by illustrating region specific sales year over year or analyse the individual productivity of a fabric. On the other hand, predictive analytics tries to predict the future based on past performance. By examining historical data, the aim is to detect patterns and relationship, to extrapolate this forward in time. For example, a commodity trader might want to foresee short term market movement, or a swimsuit manufacturer might want to predict next seasons demand of a specific model or size. Lastly, prescriptive analytics seek to minimize or maximize some objective. In the future, prescriptive analytics will perform the same tasks as predictive analytics, but it will not only make recommendations, but also independently act on them (Bellias, 2017). All these techniques leverage statistical techniques such as regression, factor analysis, multivariate statistics and knowledge of mathematics for developing equations (Dubey R. , Gunasekaran, Childe, Wanba, & Papadopoulos, 2015).

3.1.4 Industry application

These three analytical techniques did not derive from big data. However, Big data has illuminated an appetite and ability for precise, fast and actionable forward-looking insight. In recent years, big data analytics capabilities have attracted a lot of academia and management practitioners. David Kiron argued that most of the fortune 1.000 firms are investing in big data analytics related development projects (Bean, 2013). This adaption is rising quick in finance and marketing related industry fields, highlighting a predominant use of big data to understand customer intentions and behaviours (Mcafee & Brynjolfsson, 2012), whereas operations and supply chain professionals are yet exploit the true benefit of big data to improve decision-making processes (Srinivasan & Swink, 2017). Nowadays, with big data being fast and actionable, companies can create direct impact on business processes, whereas the preferable task is automatically embedded in the process itself. This can be accomplished with the help of other technologies, such as machine learning.

3.1.5 Implications on operations and supply chain management

Supply chain managers must process large quantities of data to facilitate decision making when reducing cost and increasing product availability for costumers. Historically, companies manually analysed transactional data stored in data warehouses to gain insights (Galbraith, 2014). However, companies are increasingly gaining insights from disparate data sources, developing advantages within real-time decision making. Moreover, operations and supply chain management decision are no longer constrained by internal data sources likes POS, RFID and GPS, as a vast amount of unstructured data is continuously emerging from digital clickstreams, camera and surveillance footage, social media posting, images, blog/wiki entries and forum discussions (Sanders & Ganeshan, 2015). Furthermore, supply chain activities are also supported by advanced networking technologies – sensors, tags, tracks and other smart devices, which are gathering data in real-time (Wang, Gunasekaran, Ngai, & Papadopoulos, 2016). However, various studies have proven that the best business decisions are made when the decision makers are well-equipped with data and the technical expertise to gain insights from it (Dubey & Gunasekaran, 2015). Thus, it may be argued that supply chain managers can be overwhelmed by technical features, consequently being prone to neglecting complementary assets and capabilities necessary to fully exploit the potentials.

3.2 Best practice use-cases

3.2.1 Amazon

Amazon is a market leader in collecting and using personal information from their costumers, as a mean for determining purchasing behaviour on their e-commerce platform. The company intentionally uses targeted marketing to improve user experience, customer satisfaction and build loyalty. Among other things, it enables personalised recommendations. By doing so, Amazon encourages impulsive shopping as a mean to further increase consumer spending, consummating roughly 35% of Amazon's annual sales (Wills, 2018).

Furthermore, with the acquisition of Goodreads in 2013, Amazon integrated a social networking service of approximately 25 million users to Kindle. This enabled Kindle readers to highlight words, notes and paragraphs, consequently enabling Amazon to review this data to better understand their consumer interest, thereby increasing the accuracy when recommending additional e-books.

Also, Amazon possess a patented anticipatory shipping model. It uses big data to predict which consumers are likely to purchase certain products, where and when. This allow Amazon to allot items across local distribution centres or warehouses, ready for shipping ones ordered. By continuously performing predictive analytics to increase sales and profit margins, simultaneously decreasing delivery time and holding costs.

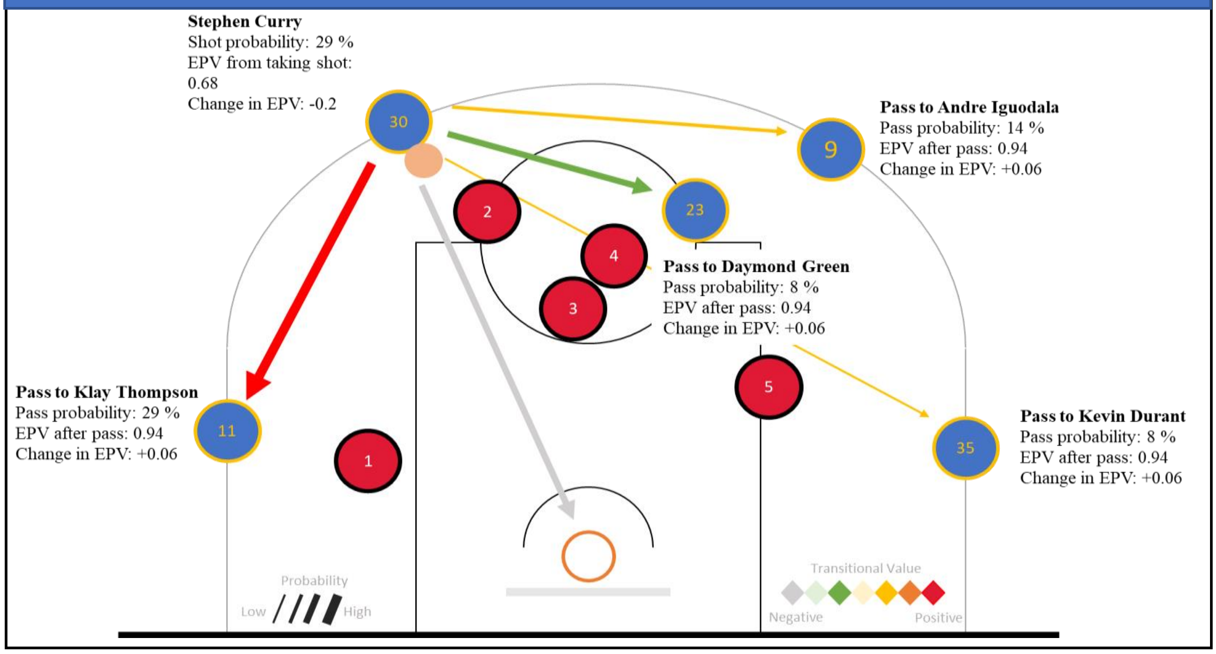
Constantly analysing data on website traffic, product availability, market pricing, trends, preferences, order history and so forth, Amazon strive to attract more customers to increase profit. Big data enable dynamic pricing, whereas, the company often sacrifice profits on best-selling items with the intention of increasing profits on less-selling items.

3.2.2 The Golden State Warriors

Data analytics is changing the NBA landscape, whereas yearly league hackathons are being held each year, to find and attract rare analytical talent. In all kinds of sports, it often comes down to gaining and identifying competitive advantage. Traditionally, NBA athletes was training hard and consistent to be quicker and stronger than their opponents. This was achieved from practicing jump shots on the quart to lifting heavier in the weight room, putting in many hours to improve their physiques. However, with increasing use of technology, data analytics are emerging as a new set of competitive capabilities for athletes and professional teams. In 2009, the league introduced a state-of-the-art video tracking technology to track player movement on the court and on the ball. With these new data sets, NBA teams are increasingly demanding data scientist knowledge, to find patterns and predicting players decision-making.

The American basketball team The Golden State Warriors, have arguably revolutionized the sport, becoming a prime example. Led by executives with limited basketball experience and Silicon Valley financiers whom bought the club in the 2010, the notion of the three-point shot was an unidentified inefficiency for the new management. With a continued ambition to win the NBA title, the team had established the so-called “death line-up” within the 2014-2015 season, with five players being threats to taking on the three-point shot. The combination of frequency and efficiency had a remarkable effect on their opponents, as they had to spread out more than before, leaving a lot of open space for the warriors to utilize. The team went on to win their first title in 40 years. In the illustration beneath, Stephen Curry for the Golden State Warriors has the ball with an Expected Possession Value of 0.68, whereas the model tries to identify what will and should happen next. Thus, by applying big data analytics, NBA players can now obtain a probability measurement on their own decision-making pattern as well as their opponents.

Big data and cartography in NBA



Head coach Steve Kerr is also continuously crediting the team’s data analysts for their success. Data is also used as a decision tool for resting to prevent injury and allow players to maximize their contribution. This has proven beneficial for the club, as players can actively contribute over longer periods, excelling their team through the league, consequently benefiting the financials of the club.

Furthermore, NBA teams are using data analysts and scientists to identify undervalued players as well as maximizing the talent of each single player. Throughout the day, teams are monitoring players with wearables to track health, injuries and fatigue values. Whereas big stars could rest in ease through their name and brand, their negotiation power is now restrained by statistics. Thus, strictly relying on continues performance to excel their career and reputation.

3.2.3 Starbucks

Starbucks does not only go through a multitude of coffee beans to stimulate consumer taste buds. The American coffee company and coffeehouse chain leverage vast amount of data to increase costumer experience and business operations. With 90 million transactions a week (Boulton, 2016), and 29.324 stores globally, almost doubled since 2008 (Starbucks, n.d.), the company is accumulating a lot of data.

Starbucks is continually incorporating big data analytics in their marketing and sales effort. One crucial aspect of Starbucks' data growth is aligned with their investments in loyalty programs and mobile payments. The mobile payment application has more than 19 million monthly users (GeekWire, 2018), whereas their loyalty program boasts more than 13 million active users with $1.2 billion cash, which is more than many banks have in deposits. Consequently, Starbucks is well-positioned to extract valuable costumer insights on purchasing behaviour. Through big bata analytics, they can find correlations between product offerings and noncompany relevant data like weather forecasts, time of hour, special promotions and so forth.

Furthermore, the company is retrieving data on coffee-buying habits from their loyalty program. This enable Starbucks to drive sales with personalized offers, such as preferable products on different hours of the day, depending on former purchases. With their stores point-of-sale system, costumers' individual preference is not tied up to a single coffee store, but their smartphone. This enable the barista to provide a set preferred and new of offerings to every single costumer.

3.3 Pros and cons

Companies have been analysing data from their own customer interactions on a smaller scale for many years, but the era of big data is still in its infancy (Ramirez, Brill, Ohlhausen, & McSweeny, 2016). Thus, extracting useful and nonobvious correlations from large data sets is a relatively new practice, with rapidly growing attention across all industries. Companies are still learning how to unlock new unprecedented insights, while still avoiding unintended or unforeseen consequences. Nonetheless, appropriately employing big data algorithms on quality data can provide market-wide benefits of more efficiently matching products and services to consumers, as well as providing numerous opportunities for improvements in society.

While recognizing potential benefits with big data, one must also address its concerns. Firstly, the collection, analysis and interpretation stages in mining large data sets may contain hidden biases, presenting considerable risk. When not properly recognized, poor data quality can lead to inaccurate decision making. Consequently, companies' risk to not provide appropriate consumer offerings or benefits. Secondly, while big data may be highly effective in showing correlations, it is axiomatic that correlation is not causation (Aldrich, 1995). Thus, with large enough data sets, companies' risk to find meaningless correlations. As a result, faulty decisions can lead to unintended consequences. These two concerns take part in a larger unified concern, where big data categorize consumers in ways that could result in excluding certain part of the population.

3.4 General future of big data

The amount of data will unquestionably continue to grow, as more connected devices and technologies will enter the market, generating larger volumes of data. Companies like Microsoft and Salesforce will to a greater extend enable non-programmers to access and modify cloud-based customer relationship management data, as well as external data sources, third-party cloud applications and on-premises enterprise resource planning systems.

Furthermore, with the rapidly increasing volume of data to prepare, analyse and group, more comprehensive conclusions can be drawn. However, the voluminous amount of data ascending in the global economy, will also make exploration of all possibilities impossible. Businesses may unintentionally exclude key insights from hypotheses data scientists did not have capacity to explore. However, Gartner estimate that more than 40% of data science task will be automated by 2020 (Panetta, 2018). They further assume that augmented analytics will identify hidden patterns, while removing personal biases made throughout the ETL process, making data insights more broadly available for all parts of the enterprise, including analysts, decision makers and operational workers.

Consumers have established an increasing awareness and value of their personal information, and their growingly concerned of how it's being accumulated by both public and private entities. Considering the general data protection regulation, data management must be grounded in ethics and trust. Corporates must both follow new legislation and question themselves whether they are doing the right thing, with an increasingly greater risk of consumer backlash and financial penalties. With the newly given access of individual data insights and guarding of intellectual property, companies must gain and maintain trust with costumers to retrieve data.

4. Big Data in Supply Chain Management

4.1 Supply chain management

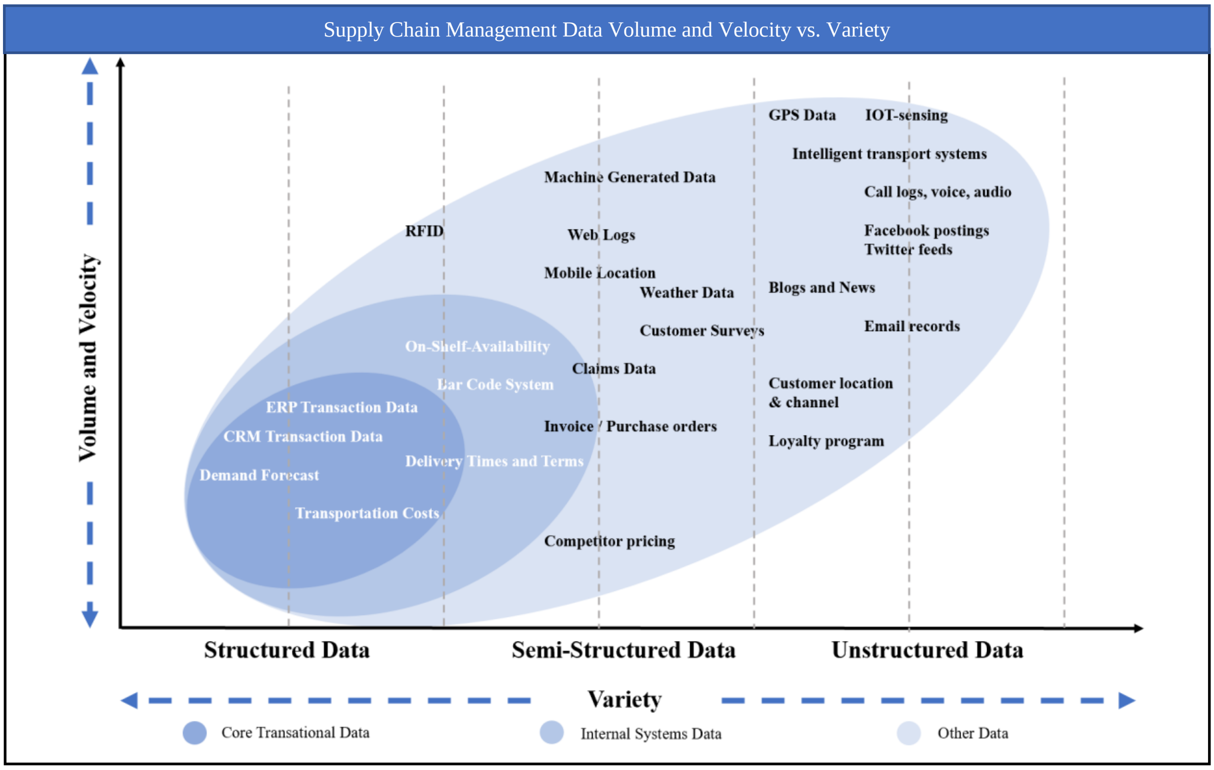
The Council of Supply Chain Management Professionals defines supply chain management as a series of key activities and processes that must be completed in an efficient and timely manner, to make products available when needed by consumers (Council of Supply Chain Management Professionals, n.d.).

Furthermore, CSCMP highlights the seven rights of fulfilment as a guide for companies to meet customer expectations with no mulligans or mistakes allowed. Through this principle, companies are inclined to execute orders according to seven specific goals to always deliver a perfect order to all customers. A perfect order is described by having the right product available for order, processing the order accurately, ship the entire order by desired means, provide the costumer with advanced shipping notification and tracking number, deliver the complete order without damaging the product, and bill the costumer correctly.

The CSCMP also state the importance of effective flow of products from point of origin to the point of consumption, to satisfy the above mentioned seven rights of fulfilment. Thus, a two-way flow of information and data should enhance the visibility of demand and enable supply chain managers to quickly detect occurring problems among its supply chain participants. This is considered a key aspect for supply chain managers decision making process. Nonetheless, products may be returned, repaired, recycled or disposed, whereas products are moving back through the supply chain.

4.2 Possibilities and limitations

Companies with the ability to harness the power of their data sources, will benefit significantly. By leveraging advanced analytics, supply chains can become more responsive, demand-driven and customer centric, as a growing number of interconnected devices, tools and applications can have deep implications on existing data sources. Nowadays, the largest data format accessible for supply chain improvement are commonly generated by unstructured data formats, as illustrated below. These are usually difficult to analyse with traditional IT tools.



Data is transforming medical practice, modernizing public policy and informing upon businesses decision making. It does also possess the ability to change existing supply chain dynamics. For example, big data analytics can utilize GPS-, weather- and traffic data to minimize delivery delays with real-time dynamic routing, as well as schedule manufacturing in line with purchasing patterns or trends. Globally, supply chains can become more proactive than reactive, responding to supply chain risk.

The explosive impact e-commerce has had on traditional brick and mortar retailer, is strongly correlated to data-driven decision-making. However, it may be argued that few companies have been able to fully apply big data analytics techniques to transform their supply chain management. Effectively analysing static and dynamic data points from the triangulation range of market, sales and social media, can provide companies the capability to predict and proactively plan supply chain activities.

The downside of the recent voluminous data burst is argued in lack of adequate knowledge. Hence, data scientist may fail to improve decision-making without domain knowledge on supply chain management activities (Waller & Fawcett, 2013). Therefore, many scholars have argued that big data analytics is overhyped. This bias is rooted in the significant awareness on big data techniques and its effectiveness, despite limited empirical research on enablers and impending barriers (Schoenherr & Speier-Pero, 2015).

4.3 Implication on various supply chain levers

Big data analytics can work across several supply chain levers. However, communicating information between organisational silos require high degree of accuracy, timeliness, consistency and completeness (Hazen, Boone, Jones-Farmer, & Ezell, 2014).

4.3.1 Marketing

Marketing has transformed costumer knowledge into agile systems where large amounts of information is sent upstream in the supply chain (Jüttner, Christopher, & Godsell, 2010). Sentiment analysis of data sources from social media, loyalty programs and mobile apps enables a greater extent of intimacy with the costumers. Similarly, recording omnichannel activates can contribute to better understanding consumers purchasing behaviour. It could be argued, that the increased information availability forces supply chain managers to increase their focus on the end-consumer, as big data technologies have made it more feasible to understand costumer data and social behaviour.

4.3.2 Sales, Inventory and Operations Planning

Planning is the understanding of real market requirements, and consolidating it into a fulfilment programme, i.e. making sure that products can be made available at the right time and place (Christopher, 2011). Furthermore, planning is one of the most data-driven processes in the supply chain. Big data can potentially redefine the planning process, using internal and external data sources. Visibility of point of sale data, production volumes and inventory data can be analysed in real-time to identify potential mismatches between supply and demand. As a result, these data sources can drive actions like changing prices, optimizing the timing of promotions or adjust content information to realign your product. One example is how meteolytix GmbH, with the help of IBM SPSS Statistics, developed an accurate sales forecast model based on weather data, historical sales and information on other contributing factors, to reduce returned goods and save costs for bakeries. By increasing forecasting accuracy, logistic capacities become more transparent, reducing product obsolescence, inventory levels and stockouts. Another forecasting example is Amazons method and system for anticipatory package shipping, where orders are packaged and distributed into the delivery network before costumers ordered them. Nowadays, companies are starting to fully comprehend big data forecasting, where the next level of sophistication will be to actively shape demand. Many retailers are already providing their consumers with individually generated product recommendations using big data analytics, inventory data management and forecasting, persuasively steering demand to available stock.

4.3.2.1 Inventory profile of the automotive supply chain

While inventory cost will vary by industry, it is suggested that the true cost of holding inventory is rarely less than 25 per cent per year of its value (Christopher, 2011). When investigating the inventory profile of the automotive supply chain, the case is a paradox where most inventory is held as finished product, i.e. when its most expensive. The true cost of the inventory to the industry is considerable. Thus, inventory management in the automobile industry is alone able to tip the balance of profit and loss. However, there is emerging a capability for proactive management, equipping automakers with the tools to rapidly respond and sense to industry changes. Essentially, automotive companies are moving from a historical point-in-time to real-time data access, pushing analytics and gaining visibility for its stakeholders. This can be highlighted in an example where automotive companies are benefitting from web interactions to forecast demand on a greater level of granularity, rapidly adapting to emerging trends on product configuration.

4.3.3 Strategic Sourcing

Strategic sourcing focuses on long-term relationship management with suppliers. It specifically aims to manage the supply base in an effective manner, by carefully selecting suppliers for long term relationships, focusing on costs, quality and delivery, by effectively manage resources and enhance supplier performance through benchmarking and continuous feedback (Talluri & Narasimhan, 2004). By leveraging supply data on procurement volumes and suppliers, companies can analyse supply processes in real-time to detect deviations from normal delivery patterns.

4.3.4 Procurement

Procurement deals with upstream activities in the supply chain. Due to a greater globalized market and cross-cultural purchasing strategies, businesses are facing greater data complexity. Within the procurement lever, there must be strong connection with internal finance reporting. To develop aggregated procurement patterns, transactional data must be made visible. Thus, companies must communicate across organisational silos.

In certain processing environment, such as pharmaceuticals, mining and chemicals, procurement can experience immense variability, even after implementing lean techniques and six sigma programs. Manufactures with a high number and complexity of productions activities are incentivised to adopt a more granular approach of diagnosing and correcting process flaws.

Furthermore, operations managers can utilize advanced analytics to investigate historical process data to identify patterns and relationships between discrete process steps and inputs to optimize factors with the greatest effect on yield.

4.3.5 Warehousing

With rapidly advancing technology, more internal and external information availability and new analytical techniques are also creating opportunities for warehousing and inventory management. Warehousing strategies are particularly benefiting from ERP data and RFID tracking. Core transactional data, internal systems and other data can be combined to create an automated sensing capability. Internet of Things and other warehouse sensor technologies can create a connectivity to drive business intelligence on material handling and packaging systems. One example is how high-rack bay warehouses possess the ability to automatically reshuffle pallets at night based on data analytics, optimizing schedules for the following day. In addition, big data is enabling chaotic storing, with the ambition to increase the efficient use of warehouse space, by minimizing distance for picking personal. Big Data also enable companies to track area-based picking performance by employers and continually optimize future allocation.

4.3.5.1 Amazon – Chaotic Storage

The multi-national e-commerce company, Amazon, reached net annual revenue of almost 178 billion U.S. dollars in 2017 (Amazon, n.d.), with nearly half of US households having an Amazon Prime subscription (Goldman, 2016). The reasoning behind the success of Prime, is heavily regarded to the instant gratification customers get when ordering a product and its quick delivery. However, Amazon's secret to moving inventory, is embedded in their Chaotic Storage Approach.

The company uses automated inventory and warehouse management systems by utilizing barcodes, RFID and other IoT sensor. Incoming products are randomly placed on available shelfing space within a given warehouse, whereas the itemized location is disregarded. The purpose is to efficiently use shelf space and minimize travel distance for staff. Through real-time data, Amazon can stimulate the increasing trend for Prime and Prime Now, removing inaccuracy from human errors, consequently benefiting lead times, reducing out-of-stock and inefficient use of storage space.

4.3.6 Transportation

Big data analytics can offer alternatives to manage and coordinate transportation in real-time. By utilizing mobile and direct sensing of delivery which are integrated with inventory. Furthermore, driving efficiencies and fuel consumption can be improved by estimating lead times when linking transportation and logistics procurement systems with traffic data, weather, marginal costs, and other big data sources. Consequently, reducing waiting time by allocating warehouses in real-time with parcel delivery expectations. Furthermore, geo-location and real-time data can enable dynamic routing of courier deliveries based on special delivery times, road regulations and traffic data.

Predictive analytics can also enable logistic providers more effective parcel delivery. By mining costumer data, companies can predict when a certain customer is more likely to be at home, potentially reducing the number of delivery attempts. In addition, companies can cut costs and carbon emissions when deciding whether transportation should be completed by train, plane, car or bike.

4.3.6.1 UPS – ORION

The courier company UPS developed ORION (On-Road Integrated Optimization and Navigation) to avoid “wrong” left turns. The system uses an algorithmic optimization, based on prescriptive analytics, to access historical and real-time data to evaluate what happened and what will happen, to guide decisions on what should happen. David Abney, CEO of UPS, said the company would save an estimated $300 million to $400 million a year thanks to ORION ((Rosenbush & Stevens, 2015).

4.3.7 Point of Sale

Retailers and manufacturers can draw insight on consumer behaviour from point-of-sale machines, providing multidimensional correlation between offerings, purchases, inventory and many other parameters. Brick and mortar, with its increasing competition from online alternatives, are increasingly paying attention to data driven optimization and how it can provide competitive advantages. These techniques can include shelf-space optimization and mark-down pricing. For example, brick and mortar retailer may harness data analytics to optimize product selection for high valuation store location, like end-of-aisle. Data derived from point of sale can be utilized to better understand consumers responding to product pricing, thus increasing their ability to manage inventory levels in real-time. Product codes can be used to visualize consumer trends and stock availability real-time and during clustered time periods.

As RFID tags are still considered too expensive to be applied for single grocery items, retailers are monitoring sales activity from out-of-stock indicators. Point-of-sale data can develop market trends and identify when certain products are assumed out-of-stock. Combined with machine learning capabilities, systems can automatically develop stocking orders, whereas point-of-sale driven data management can provide companies a better understanding of the optimal product mix for maximized profit.

4.4 Future big data implications on supply chain

A research by SCM World, showed that 64% of supply chain executives consider big data as a disruptive and important technology, setting the foundation for long-term change management. However, only 17% percent having reported on big data implementation in one or more supply chain functions (Accenture, 2014). With continuously increasing volumes of unstructured- and structured data sources, big data technologies can contribute to more visible supply chains, providing deeper insights. Thus, it may be argued that more supply chain executives will enlarge their big data analytic capabilities in the coming years.

In a time where costumer expectations are increasing, it is key for companies to offer the right product, at the right time and place, to the right costumer to gain or retain loyalty and satisfaction. With big data capabilities increasing, businesses can leverage an increasing 360-

degree view of costumers, using predictive techniques to gain valuable insight on customer needs, preferences and customized brand experience.

Big data is also predicted to further increase supply chain efficiencies, reducing costs and lead times with spend analytics. According to Accenture, embedding big data techniques in operations can improve supply chain efficiencies by 2.6x (Accenture, 2014). In addition, greater visibility and predictability in supply chain operations can help reducing supply chain risks. For example, prescriptive maintenance will enable companies to move away from compliance with time-based- and condition-based maintenance. Where predictive maintenance has taken condition-based maintenance further, prescriptive maintenance will not only provide recommendations, but independently act on them. Hence, companies may evolve their maintenance systems from being simply efficient to becoming strategic. However, these systems must be cognitive, with the ability to integrate operations data with other data, such as weather, quality, warranty and engineering data, to gain insights on how the entire company operates.

Aligned with rapidly changing customer expectations, agility and speed are becoming more crucial. Hence, to meet everchanging fulfilment objectives, the ability to be flexible and rapid is one of the most important drivers for competitive advantages. Accenture estimated that big data analytics in operations can provide a 4.25x improvement in order-to-cycle times (Accenture, 2014).

5. Qualitative Interviews

To further evaluate the implications of big data and analytics, three qualitative interviews were arranged with experienced data scientist working within consulting and banking. The combined outcome from the interviews and secondary data sources from literature, will be used to derive theories on current and future applications, implications and developments.

Qualitative Study: Analysis of Expert Interview

Assertions	Opinions	Comments
Insights from Big data and analytics will yield benefit.	Strongly Disagree — Strongly Agree	General Agreement
Big data is applicable to all industries.	Strongly Disagree — Strongly Agree	General Agreement
Deploying Big data strategies are always the same.	Strongly Disagree — Strongly Agree	Slight Disagreement
Augmented Analytics is future of big data and analytics	Strongly Disagree — Strongly Agree	Slight Disagreement
Clients know what kind of insights they're dataset will aggregate.	Strongly Disagree — Strongly Agree	Large Disagreement (Project Oriented)
Data regulations will have an impact on data aggregation and handling	Strongly Disagree — Strongly Agree	General Agreement

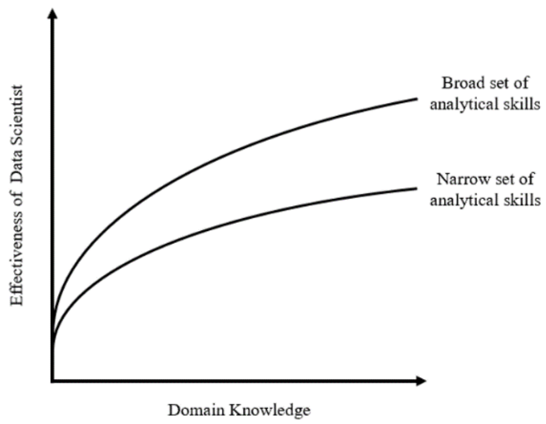
5.1 Impact on Business Revenue

Big data does not directly affect business revenue, as the data itself does not yield benefits by itself. However, by combining big data and analytics, companies can create valuable business insights, which were not necessarily visible. To unravel these insights, adequate candidates can utilize data from various sources to test a hypothesis. The hypothesis is then used to verify correlations, consequently increasing the understanding of a specific business matter, whether its maintenance in manufacturing processes or further developing an understanding of the end consumer. Nowadays, companies can easily retrieve large proportions of data, whereas the data does not lie. However, companies must understand that correlation does not mean causation, thus further increasing the importance of quality datasets.

5.2 The necessity of data science and domain knowledge

One of the limitations within big data generated strategies is the lack of qualified knowledge. Data science is not agnostic to domain knowledge, creating the necessity for a combined skillset. However, acquiring strong analytic capabilities takes a lot of time. Likewise, deep domain knowledge on supply chain management may only be acquired through accumulated experience, only accessed by motivation and well-invested time. Consequently, candidates with both analytic skills and domain knowledge are extremely rare to come by.

However, an increasing focus on data- processing and aggregation technologies, as well as augmented analytics may pave the way for big data in supply chain management. With increasing focus on data pipeline aggregation, machine learning and visual tools, supply chain professionals with proper domain knowledge can obtain a broader analytical mindset, as data science and domain knowledge thrive in combination with each other. Thus, making data insights more accessible and affordable.



5.3 Big data strategies and scalability

Even though big data can provide unprecedented value with its insights, the way data analytics is deployed may vary from one company to another, as well as in between industries. For example, technological companies and platform providers often possess a lot of data on user behaviour. In this industry, data sets are extremely large and can scale to several millions every month. Typically, these companies aim to utilize big data and analytics to better understand the end-user. As service providers increasingly understand their consumers, personalized offerings are becoming a norm. Utilizing the understanding of end-consumers can increase customers satisfaction and loyalty, but it’s also raising costumer expectations.

Another limitation for big data and analytics is the use case. Many companies and executives are seeking a specific explanation or answer, even though the hypothesis tested may not create any actionable insights. Thus, data scientists are heavily dependent on the size and quality of the dataset. A correlation might not hold any value, whereas supply chain managers must carefully navigate and understand the data availability and origin, before blindly making decisions based on evidence in data.

Furthermore, with a growing access and need for data in decision making, companies must develop comprehensive scalable data platforms. Even though big data analytics can provide unprecedented insights, companies must thoroughly evaluate the scalability of big data strategies with existing strategies, as they may also face intricate challenges from regulations, scale and other technological and human complexities.

5.4 Impact of data regulations

Big data comprises voluminous amount of personal data, thus involving strong requirements on data-privacy and protection. These requirements are further empowered by the European General Data Protection Regulation (GDPR), whereas companies must oversee compliance and security than before. These legislations will create a larger reliance on real-time analytics with faster turnaround. Furthermore, the ability to utilize social media platforms for building customer loyalty and increasing engagement, will prove more difficult with an increasing introduction of new privacy tools.

These regulations may at first seem like a burden for companies to overcome, but an increasing amount of data breaches have left end-consumers increasingly concerned. Consumers may thus be reluctant to provide personal information, in fear of being comprised. Hence, companies with the ability to upheld compliance, can achieve competitive advantage through long-term trust.

6. Conclusion

The goal of this master thesis was to understand big data as a technological enabler for new operating business models, by investigating its limitations and implications on general use-cases as well as supply chain management levers. With seemingly endless application scenarios of big data, it was crucial to verify secondary data sources with qualitative interviews, to draw a conclusion. This achievement will be summarized and critically reflected upon in the following paragraph.

I can hereby conclude that appropriate use of big data in supply chain management can enhance supply chain levers, by increasing automation efficiencies and facilitate more proactive and effective business decision making processes. In a future end-to-end digitalised supply chain, big data can enable self-sufficient decision-making by machines, with an ability to continuously sense and quickly respond to shifting patterns in supply and demand. As of now, some activities proposed have a rather disruptive approach to current supply chain levers, where holistic changes and large investments must be made to cohere with existing strategies. However, in other cases, big data can offer substantial improvements on efficiency with minor modifications. To succeed with big data, companies must view information as a strategic asset, as it possesses economic value with a capitalizing potential. Even though big data can provide unprecedented insights and opportunities across a multitude of industries, it also raises concerns and questions to be addressed. With everchanging data analytics techniques and tools, stricter governmental regulation on data scrutiny and a surge for adequate talent, companies must evaluate the scalability and profitability of big data strategies with existing strategies.

Bibliography

The Accenture. (2014). *Big Data Analytics in Supply Chain: Hype or Here to Stay?* Accenture.

Aldrich, J. (1995). Statistical Science. *Correlations Genuine and Spurious in Pearson and Yule*, 364-376.

Amazon. (n.d.). *Net sales revenue of Amazon from 2004 to 2017 (in billion U.S. dollars). .* Retrieved from In Statista - The Statistics Portal: https://fesrvsd.fe.unl.pt:2351/statistics/266282/annual-net-revenue-of-amazoncom/.

Bauer, C. (2015). *Industry 4.0: How to navigate digitalization of the manufacturing sector.* Munich: McKinsey Digital.

Bean, R. (2013, January 28). Organizational Alignment is Key to Big Data Success. (D. Kiron, Interviewer)

Bellias, M. (2017, March 14). *IBM*. Retrieved from The evolution of maintenance: https://www.ibm.com/blogs/internet-of-things/maintenance-evolution-prescriptive/

Boulton, C. (2016, April 1). *CIO*. Retrieved from Starbucks' CTO brews personalized experiences: https://www.cio.com/article/3050920/analytics/starbucks-cto-brews-personalized-experiences.html

Buneman, P. (1997). *Semistructured Data.* Retrieved from http://sites.fas.harvard.edu/~cs265/papers/buneman-1997.pdf

Cheng, S., Zhang, Q., & Qin, Q. (2016). Big data analytics with swarm intelligence, Industrial Management & Data Systems.

Christopher, M. (2011). *Logistics & Supply Chain Management.* Pearson.

Council of Supply Chain Management Professionals. (n.d.). *Supply Chain Management Concepts*. Retrieved from CSCMP: https://cscmp.org/CSCMP/Develop/Starting_Your_SCM_Career/SCM_Concepts/CSCMP/Develop/Starting_Your_Career/Supply_Chain_Management_Concepts.aspx?hkey=96af0d8b-21ad-4bca-b7d1-956a25ced524

Dubey, R., & Gunasekaran, A. (2015). Supply chain talent: the missing link in supply chain strategy. *Industrial and Commercial Training, Vol. 47 Issue: 5*, 257-264.

Dubey, R., Gunasekaran, A., Childe, S. J., Wanba, S. F., & Papadopoulos, T. (2015). The impact of big data on world-class sustainable manufacturing. *The International Journal of Advanced Manufacturing Technology*, 1-15.

Economist, T. (2010, February 27). *Data, data everywhere.* Retrieved from https://www.emc.com/collateral/analyst-reports/ar-the-economist-data-data-everywhere.pdf

Franks, B. (2012). *Taming The Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics.* Wiley.

Galbraith, J. R. (2014, 2 13). Retrieved November 2019, from file:///C:/Users/Henrik/Downloads/8856-Article%20Text-38451-4-10-20140522%20(1).pdf

Gartner, I. (n.d.). *IT Glossary*. Retrieved from Big Data: https://www.gartner.com/it-glossary/big-data/

GeekWire. (2018, December 20). *Number of Starbucks monthly mobile app users in the United States as of April 2016 (in millions). .* Retrieved from In Statista - The Statistics Portal. : https://fesrvsd.fe.unl.pt:2351/statistics/629431/starbucks-mobile-app-users/

Goldman, D. (2016, April 28). *Amazon shares soar on fourth straight profitable quarter.* Retrieved from CNN Business: https://money.cnn.com/2016/04/28/technology/amazon-earnings/

Hazen, B. T., Boone, C. A., Jones-Farmer, L., & Ezell, J. D. (2014). Data Quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An Introduction to the Problem and Suggestions for Research and Applications. *International Journal of Production Economics*.

Inmon, W. H., & Krishnan, K. (2010). *Building the Unstructured Data Warehouse: Architecture, Analysis, and Design.* Technics Publications.

Jacobsen, R. (2013, April 24). *2.5 quintillion bytes of data created every day. How does CPG & Retail manage it?* Retrieved from IBM Industry Insight: https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/

Jüttner, U., Christopher, M., & Godsell, J. (2010). A Strategic Framework for Integrating Marketing and Supply Chain Strategies. *The International Journal of Logistics Management*.

King, J. H., & Richards, N. M. (2014, 3 28). *Forbes*. Retrieved from What's Up With Big Data Ethics?: https://www.forbes.com/sites/oreillymedia/2014/03/28/whats-up-with-big-data-ethics/#5620f61e3591

Laney, D. (2001, February 6). *3D Data Management Controlling Data Volume, Velocity and Variety.* Retrieved from https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Mcafee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*.

Neuman, W. (2003). Social Research Methods: Qualitative and Quantitative Approaches.

Oracle. (n.d.). *What is Big Data?* Retrieved from https://www.oracle.com/big-data/guide/what-is-big-data.html

Panetta, K. (2018, October 15). *Gartner Top 10 Strategic Technology Trends for 2019*. Retrieved from Gartner: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/

Ramirez, E., Brill, J., Ohlhausen, M. K., & McSweeny, T. (2016, January). *Big Data - A Tool for Inclusion or Inclusion.* Retrieved from https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf

Sanders, N. R., & Ganeshan, R. (2015). Special Issue of Production and Operations Management on "Big Data in Supply Chain Management".

SAS. (n.d.). *ETL*. Retrieved from What it is and why it matters: https://www.sas.com/en_us/insights/data-management/what-is-etl.html

Schoenherr, T., & Speier-Pero, C. (2015). Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential. *Journal of Business Logistics*.

Srinivasan, R., & Swink, M. (2017). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective.

Starbucks. (n.d.). *Number of international and United States Starbucks stores from 2005 to 2018.* Retrieved from In Statista - The Statistics Portal: https://fesrvsd.fe.unl.pt:2058/statistics/218366/number-of-international-and-us-starbucks

Talluri, S., & Narasimhan, R. (2004). A Methodology for Strategic Sourcing. *European Journal of Operational Research*.

Thomas H. Davenport. (2013). *Deloitte.* Retrieved from The Analytics Advantage, We're just getting started: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Deloitte-Analytics/dttl-analytics-analytics-advantage-report-061913.pdf

Waller, M. A., & Fawcett, S. E. (2013). A Data Scientist: Big Data, Predictive Analytics, and Theory Development in the Era of a Maker Movement Supply Chain. *Journal of Business Logistics*.

Wang, G., Gunasekaran, A., Ngai, E., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: certain investigations for research and applications. *International Journal of Production Economics, Vol. 176*, 98-110.

Wills, J. (2018, October 20). *Investopedia*. Retrieved from 7 Ways Amazon Uses Big Data to Stalk You (AMZN): https://www.investopedia.com/articles/insights/090716/7-ways-amazon-uses-big-data-stalk-you-amzn.asp