FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

THE RELATIONSHIP BETWEEN BIG DATA ANALYTICS AND OPERATIONS RESEARCH

LITERATURE STUDY, APPLICATIONS AND RESEARCH OPPORTUNITIES

Word count: 34.052

Nathan De Coninck
Student number: 01202143

Supervisor: Prof. dr. Broos Maenhout

Master's Dissertation submitted to obtain the degree of:

Master of Science in Business Engineering

Academic year: 2016 - 2017

UNIVERSITEIT
GENT

FACULTEIT ECONOMIE EN BEDRIJFSKUNDE

THE RELATIONSHIP BETWEEN BIG DATA ANALYTICS AND OPERATIONS RESEARCH

LITERATURE STUDY, APPLICATIONS AND RESEARCH OPPORTUNITIES

Word count: 34.052

Nathan De Coninck
Student number: 01202143

Supervisor: Prof. dr. Broos Maenhout

Master's Dissertation submitted to obtain the degree of:

Master of Science in Business Engineering

Academic year: 2016 - 2017

UNIVERSITEIT GENT

Foreword

For you lies the master's dissertation concerning the relationship between Big Data Analytics and Operations Research. This relationship is assessed in literature on what has already been written, which applications are available and what opportunities exist in future research. This dissertation is written as a final stage in order to obtain the degree of Master of Science in Business Engineering at the University of Ghent. The establishment of this dissertation was an expiring, insightful and challenging but captivating experience that has altered the way I think and approach things.

First and foremost, I would like to thank my supervisor Prof. dr. Broos Maenhout for his help and guidance during the past two years. His critical view and constructive remarks lifted this dissertation to a higher level and ensured certain insights for my sake. The communication went smoothly, he provided me with relevant and up-to-date literature and the door was always open. I would also like to acknowledge the support of my girlfriend Veerle. I won't forget the motivational speeches and encouraging words during the progression of our work. A special mention for Koen Cooreman, advisor Information Security at FOD P&O in Belgium, is also in place. He revised this dissertation multiple times and provided valuable and extensive feedback. Finally, I want to give my special thanks to my parents, who gave me the opportunity to start and successfully finish my studies at the University of Ghent.

Nathan De Coninck

Ghent, Belgium

May 11th, 2017

Table of Contents

List of Figures

List of Tables

1. Introduction

"Consumer data will be the biggest differentiator in the next two or three years. Whoever unlocks the reams of data and uses it strategically will win." – Angela Ahrendts, Senior Vice President of retail at Apple

"Information is the oil of the 21st century, and analytics is the combustion engine" – Peter Sondergaard, Senior Vice President at Gartner Research

"Hiding within those mounds of data is knowledge that could change the life of a patient, or change the world" – Atul Butte, professor at University of California

"As business leaders, we need to understand that the lack of data is not the issue. Most businesses have more than enough data to use constructively; we just don't know how to use it. The reality is that most businesses are already data rich, but insight poor." – Bernard Marr, bestselling Big Data author and guru

"Big Data is at the foundation of all of the megatrends that are happening today, from social to mobile to cloud to gaming." – Chris Lynch, renowned published writer

Above quotations show the relevance of Big Data and the need for a clear definition, framework or demarcation. Big Data nowadays is ubiquitous (Assunção, Calheiros, Bianchi, Netto, & Buyya, 2014; Bughin, 2016). In combination with Operations Research, great opportunities can be captured. Operations Research has had an enormous impact on organizations and the corporate world in the last 70 years (Poppelaars, 2013; Ranyard, Fildes, & Hu, 2015). Recent developments of new algorithms, informatics and data management have recently created an environment wherein the domain of Operations Research/Management Sciences (OR/MS) will have even a bigger impact. By combining predictive analytics, like data mining and statistics, with prescriptive analytics like optimization methods, one is capable to develop all kinds of new applications inside an organization (Mazzocchi, 2015). Ergo, taking qualitative decisions is not solely the result of data analysis. An efficient Decision Support System (DSS) is indispensable. A decision is supported by (big) data, objectives, decision criteria and a decision-supporting model that gives insight all various alternatives and restrictions of the problem.

Loads of data surround us with or without our knowledge. Since the digitalization in business processes, a majority of companies and governments all around the world are exposed to large amounts of data, without knowing how to handle them correctly (Meyer, McGuire, Masri, & Shaikh, 2013). In scientific literature, there is no extensive overview of the most prominent current trends and applications of Operations Research domains that make use of Big Data. Numerous published articles each describe an in-depth analysis of one particular application inside a certain business domain. Now and then, an article tries to outline an overview of some trends within a business domain, but the scope is almost always very narrow. Recently, there has been a call by Elsevier to gather articles with a subject regarding "*Emerging trends, issues and challenges in Internet of Things, Big Data and Cloud Computing*"[1]. Clearly, there is a high need for an extensive overview of application and trends of Big Data, preferably concretized in Operations Research domains. This is where this master's dissertation comes in.

After this introduction, some important and highly relevant concepts will be explained in *section 2*. In *section 2.1 and 2.2*, Big Data Analytics and Operations Research are shortly delineated separately at first and then together as well in *section 2.3*. During the last few years, rapid changes have highly influenced the development of both domains. Therefore, the evolution towards a more and more integrated approach of both business domains is depicted as the origin of this closer approach is key. After this introductory concept overview, the research objectives are defined in *section 3*. The main objective of this dissertation is to give a comprehensive overview of the most prominent trends and applications in several OR related business domains. This section will emphasize the importance of this objective. After this, the methodology is described in *section 4*.

Section 2 already attempts to prove the importance and the call for a concise and integral overview of these trends and applications, via the considered concepts. This objective is elaborated in the body of the text, namely *section 5*, along with the research opportunities and threats of these ongoing trends. Several important business domains are scrutinized and a clear overview of the most prominent trends and applications is given. In *section 5.1* we dive into the use of Big Data Analytics in the marketing area. With its wide application field and abundance of data from all kinds of sources, it is self-evident that Big Data Analytics can create value for marketing when it comes to customer approach. After this, the prominent applications of Big Data in healthcare are discussed in *section 5.2*. The rise of Big Data Analytics is of vital importance in the healthcare environment, as it can save lives when implemented or applied correctly. In this area, applications have a wide application, going from patient monitoring, to personnel scheduling and even fraud detection. *Section 5.3* then delineates the use of Big Data Analytics in operations and supply chain management (O/SCM). As this is also a very large and very diverse area, many applications of Big Data are possible. We tried to focus on the most prominent applications and provided use cases accordingly. In *section 5.4*, we also take a look at the public services and how they already benefit from integrating Big Data Analytics in their daily operations. Exuberant volumes of data are in the possession of public authorities, but they often stay unused. This potential goldmine can be broached by the rise of Big Data Analytics and the further adoption in public services. Of course, not all that glitters is gold, so at

[1] https://www.journals.elsevier.com/future-generation-computer-systems/call-for-papers/special-issue-on-emerging-trends-issues-and-challenges-in-in

the end of the body some major concerns and threats are outlined and brought in perspective, in *section 6*. Concerns related to privacy, security and ethics are extensively addressed as these will be the main challenges in the future. They will determine which direction Big Data Analytics will take in OR. Challenges related to technology, data characteristics and the nature of OR are also briefly delineated. This dissertation ends with a concise summary of the research, indicates what the contribution is to the scientific community and what the limitations are of this research.

2. Concepts

In this section, we will explain a few concepts that are crucial in order to fully capture the essence of this dissertation. In *section 2.1*, the differences and similarities of Big Data and Analytics are explained, resulting in a clear explanation of Big Data Analytics in general. Then, in *section 2.2*, Operations Research is shortly delineated, focusing on the essential components that made OR to the extensive research domain it is today. After these 2 sections, Big Data Analytics in Operations Research is discussed in *section 2.3,* since this interrelatedness is the breeding ground for all applications discussed in *section 5*, later on in the text. Here, we'll explain how these former different domains have grown towards each other and how they can benefit both from an integrated approach. The methodology of how Big Data is used in Operations Research is described as well, as this is essential in order to understand how these two domains can create a synergy towards better results*. Section 2.4* gives a deeper understanding of how decisions are finalized after the business model is made and analyzed with data. A well-structured and up-to-date decision support system is indispensable for every organization in the 21st century, as this section will prove. But first things first, the driver of all this: Big Data Analytics.

2.1 Big Data Analytics

2.1.1 Big Data

Big Data and Big Data Analytics are two very similar yet different things. **Big Data** is a concept that does ring a bell with many professionals nowadays and strongly gains popularity and adoption (Hashem et al., 2015; Konstantinova, 2014). It is often defined on the basis of the **3V-model**, as *figure 1* shows (Fogelman-Soulié & Lu, 2016; Franková, Drahosova, & Balco, 2016). The continuously growing amount of data (*Volume*) which are not only present on the internet, but also internally generated by companies and individuals, appear in all kinds of different formats (*Variety*) and are generated and stored more rapidly (*Velocity*). The volume of data has grown from terabytes (10^{18} bytes) of information towards zettabytes (10^{21} bytes). A once very structured dataset has now become a combination of structured and unstructured data in all kinds of formats and the velocity of data exposure, data acquisition and data processing has gone from batching towards streaming. In se, the amount, speed and diversity of data define the concept Big Data, but there are other factors that must be considered. According to Katal, Wazid, and Goudar (2013), there is a certain variability (not the same as variety) present in the data and in the quality of them. Several sources may contradict each other and this enhances the complexity of analyses. Some sources are more reliable than others and this encompasses a great danger. A strict analysis therefore imposes itself before taking any conclusions from the self-evident sources.

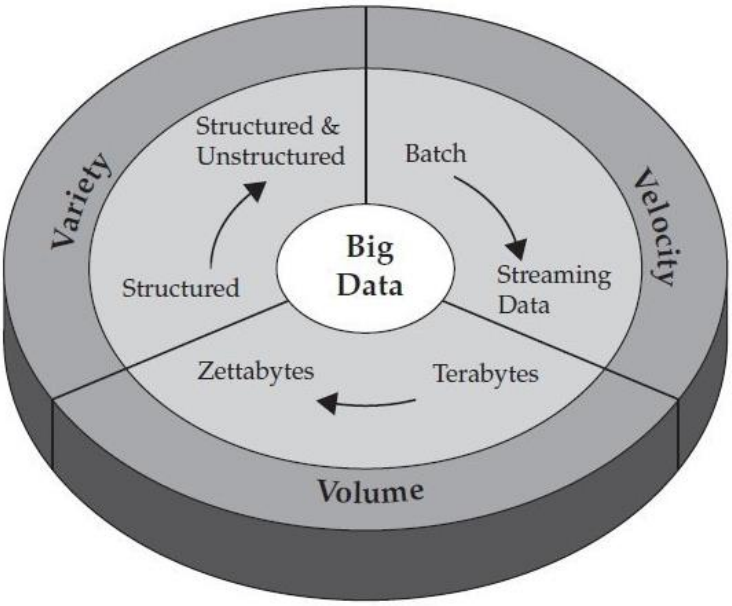

Figure 1: The 3V-Model

The continuously evolving data have a great impact on companies and their way of doing business. The mass of available and often unstructured data has a priceless value, provided that it is processed and analyzed correctly, after which decision makers can elaborate appropriate business actions. Here, a first problem comes into the picture. Nowadays, many companies collect lots of data, but they are struggling to process them correctly, let alone that they can analyze them correctly or to connect the right business actions to it (Desouza, 2014; Vidgen, Shaw, & Grant, 2017). This brings us to the **three defining aspects of Big Data**: volume, variety and velocity.

Volume

If we first dive into the concept of **volume**, it is safe to say that the amount of data generated each day has grown astonishingly hard (Gantz & Reinsel, 2012; Chen, Chiang, & Storey, 2012; Bello-Orgaz, Jung, & Camacho, 2015). Until 2003, 5 exabytes of data in total were created worldwide. In 2012, the total amount of data ever generated had expanded to 2.72 zettabytes. It is predicted to double every two years now the expansion is in its full growth stage – following Moore's law – after which the growth will decline to a rate of 2.3 times the original, reaching a staggering amount of 180 zettabytes of data by 2025, according the IDC (Press, 2016). This increase in volume of data is also reflected in the financial growth expectations of companies dealing with Big Data and business analytics software. Gartner (2016) predicts that Big Data and business analytics worldwide revenues will grow from nearly $122 billion in 2015 to more than $187 billion by 2019, an increase of more than 50% over a five-year period (Columbus, 2016). Along with the growth of produced data come new technologies to cope with these data and to translate correlations, relations and associations into business insights. These are the basis for concrete business actions and that's what it's all about. The term 'big' in Big Data is not per se defined as a minimum volume size a dataset needs to meet. It is big in all its facets. The data are numerous, the technological options to handle them cannot be counted on one hand, open sources have immeasurable reaches… Big Data is more than just a large amount of data gathered somewhere.

Variety

The second V of the 3V-model corresponds to the diverse representations and appearances of data: **variety**. Data are generated in many different ways and are stored accordingly. To address these different forms and representations of data, a strong computational and comprehensive technology is required, which makes the data Big Data. Data do not have a fixed structure and are often not structurally ordered nor are they ready for processing (O'Reilly Media, 2015). Data formats range from highly structured (relational database data), over semi-structured (social media data, sensor data, email, web logs…) to unstructured (video, still images, clicks, audio…). Variability is often seen as an extension of the second V, to emphasize on the semantics, or the variability of meaning in language and communication protocols (Emani, Cullot, & Nicolle, 2015).

Velocity

The V that makes the 3V-model complete is **velocity**. It goes without saying that this is an essential part of Big Data as well, since time is money in the business world. Data is produced real time and needs to be processed as fast as possible in order to be competitive sometimes. Hence, it is not only the velocity of the incoming data that is important, but the speed of the feedback loop as well. The time from data instream towards decision needs to be small. Preferably as small as possible, but hasty decisions are odious as well.

2.1.2 Analytics

If we broaden our scope from Big Data towards **Big Data Analytics**, the main difference is the analytical part of doing business. Big Data is mainly used to analyze insights which can lead to certain strategic business moves and better decisions. Big Data Analytics is one step further. It involves automating insights into a certain dataset as well as it supposes the usage of queries and data aggregation procedures (Monnappa, 2016).

Four key drivers

According to Liberatore and Luo (2010), Big Data Analytics has **four key drivers**, as depicted on *figure 2* (Liberatore & Luo, 2010). These four drivers each have a strong independent effect on the outcome of the analysis, but when taken into account together and through their interactions, they are a powerful force leading to the growth on the analytic organization.

The **data** aspect is self-evident and is already depicted in the paragraphs above, as this is the handhold between Big Data and Analytics. The **people** that are needed to bring analytics to a good end need to be highly technology-minded and need a strong analytical mindset in order to cope with all the data. They need to automate getting insights out of the data via a certain software. There are many simple statistical and optimization tools out there, but in order to be doing some "real" Big Data Analytics, a more advanced **software** package is often needed. The fourth key driver emphasizes the need for a **process** orientation to better understand the tasks that comprise the firms' businesses. This process orientation has to align the analysis tasks with the corporate objectives. This is important to make progress and not continuously being stuck in the same endless cycle of executing irrelevant tasks. By keeping an eye on the objectives, the three E's can be realized: Efficiency, Effectiveness and Economy.

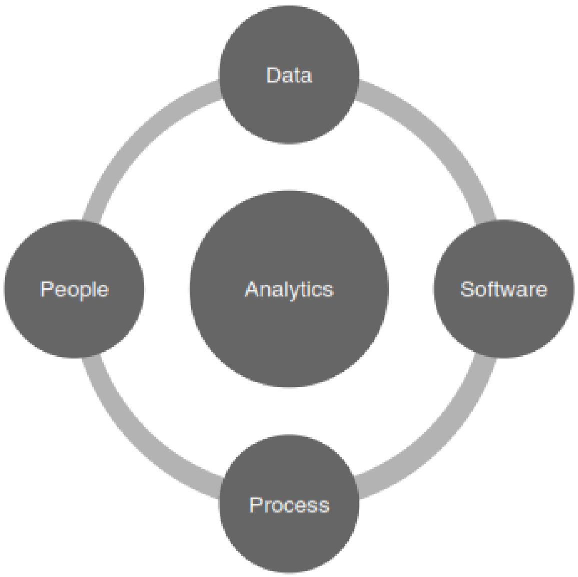

Figure 2: Four Key Drivers of Big Data Analytics

Descriptive, Predictive & Prescriptive Analytics

The application of analytics can be divided into three main categories, namely descriptive, predictive and prescriptive analytics. **Descriptive analytics** involves using advanced techniques to locate relevant data and identify remarkable patterns in order to better describe and understand what is going on with the subjects in the dataset and hence in real-life. Data mining, the computational process of discovering patterns in large datasets involving methods at the intersection of artificial intelligence, machine learning, statistics and database systems, is accommodated in this category (Sumathi & Sivanandam, 2006). A descriptive model thus describes what has happened, but a description on its own is never enough for decision makers (Barceló, 2015). Descriptive models can give a clear explanation why things behaved the way they do and why certain events occurred, but all this already is past perfect. Companies can watch back to the past and see what happened, maybe even what is the cause, but what is important is the future and how they should behave in the future. This calls for predictive models. **Predictive analytics** is most often seen as a subset of data science (Waller & Fawcett, 2013; Hazen, Boone, Ezell, & Jones-Farmer, 2014). Liu and Yang (2017) formalize how a predictive OR model is made self-organizing via Big Data. It makes use of data, statistical algorithms and various machine learning techniques to identify the likelihood of future outcomes based on historical data (Bose, 2009). The built model thus predicts (hence the name) what is likely to happen, based on the available data. Therefore, a rich and extensive dataset is key. The amount of data available is not the problem, the richness of the data however is often questionable. This is most certainly required when people want to perform **prescriptive analytics.** When executed right, this application of mathematical and computational algorithms enables decision-makers to not only look into the future of their own processes and opportunities, but it even presents the best course of action to take to gain advantage of the foresight, based on the data. The requirements for an accurate and reliable prescriptive analytics outcome are hybrid data, integrated predictions and prescriptions, taking into account side effects, adaptive algorithms and a clear feedback mechanism (Basu, 2013). The ultimate example of a prescriptive model is the decision support system (see *section 2.4 Decision Support System*). *Figure 3* displays the three stages in terms of value/intelligence and difficulty (Gartner, 2015).

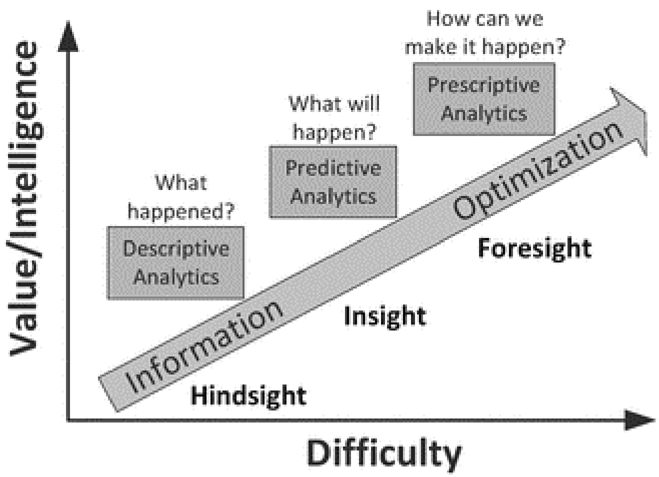

Figure 3: The Three Phases of Analytics

The relatively new field of analytics matures and gravitates from a primary focus on the statistical and econometric models of descriptive and predictive analytics to prescriptive analytics, with its focus on Operations Research and Management Science (OR/MS) optimization models and decision support systems (Levasseur R. E., 2015).

2.1.3 Big Data Analytics

Big Data Analytics thus is more than just the ubiquity of data. It is the combination of the volume, variety and velocity of data and the three separate phases or approaches that makes people say Big Data Analytics. Other authors mainly focus on three other V's, namely Value, Veracity and Visualization (Hitzler & Janowicz, 2013; Gandomi & Haider, 2015). Liberatore and Luo (2010) say data analytics is a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving. This gives a clear and concise idea to conclude with Big Data Analytics, as you can see on *figure 4* (Liberatore & Luo, 2010). The final action is the most important and distinctive decision for managers, but without a rich dataset and a huge variety and volume of data, operational actions are limited or less-informed.

The richness of the dataset is determined by the collection, extraction and manipulation of **data**. Proficient data scientists need to do this job neatly, as this determines how strong the subsequent analysis can be. This **analysis** exists of three phases. Primarily, data are visualized and analyzed using charts, dashboards and interactive tables as this can be understood by more people and tends to show more than plain numbers. According to Shneiderman (2014), more people are able to see the big picture when there is a clear data visualization. Secondly, the predictive model tries to estimate trends, relationships and classifications based on the data input. Finally, an optimization model seeks to find the optimal solution, subject to a set of assumptions and constraints. This analysis makes sure that clear **insights** can come to light. During the insight phase, managers have to think about the past, the current situation and the future. Based on all the foregoing, specific **actions** need to be taken. These actions range from operational decisions to process changes all the way up to strategic formulations.

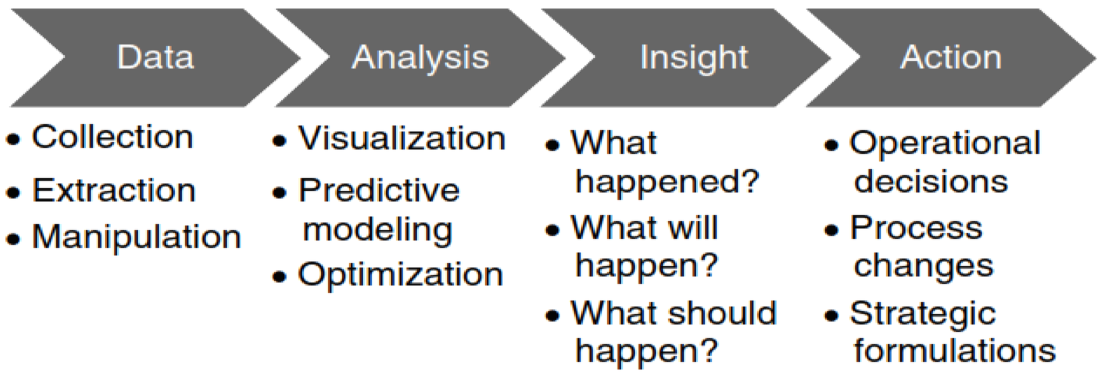

Figure 4: Big Data Analytics Process Steps

Although there is some indisputable evidence of the beneficial impact of Big Data Analytics on the decision-making process (*see section 2.4 Decision Support System*), some authors (Barton & Court, 2012;

Waller & Fawcett, 2013) say Big Data (Analytics) is simply the next buzzword of the day, nothing more. This is somewhat understandable given the rapidly evolving definition of the concept (Gandomi & Haider, 2015) and the fact that there exist a lot of different definitions. People tend to misuse the whole concept of Big Data and try to flow on its popularity. In 2013, the National Science Federation had a call for proposals on Big Data definitions (McLay, 2013). Their summary of a concise definition is the following:

> *"Big Data refers to large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future."*

Gudivada, Baeza-Yates, and Raghavan (2015) define Big Data as

> *"data too large and complex to capture, process and analyze using current computing infrastructure. It is defined by five V's, Volume, Velocity, Variety, Veracity and Value."*

Despite that most definitions focus mainly on the data characteristics alone and how they are acquired, it strongly implies the importance of the data, as every great action is built on great data. But data without a correct model or analysis are just plane data, a silent gold mine that doesn't say anything. Therefore, the analysis is a very important phase and needs to be executed with the organizational objectives in mind. How to build a correct model in order to analyze the data correctly, is explained in the following section.

2.2 Operations Research

Operations Research (OR) in general is not the main focus of this research paper, so this section will rather give a short definition and demarcation of this discipline. Hence, the reader gets a better insight of how Operations Research is approached in the remainder of the text. Where needed, additional sources are denoted for the interested reader.

Operations Research involves, as the name implies, research on operations. Thus, it is applied to problems that concern how to conduct and coordinate the activities (i.e. the operations) of an organization (Hillier & Lieberman, 2015). There is a wide application area of Operations Research. To name a few industries where OR is extensively used: marketing, healthcare, operations and supply chain management, public services, financial services, energy sector, environmental sector etc. This list can be as long as you want, but the first four business domains will be further elaborated in *section 5. Application & Trends*. According to some authors (Du, Hu, & Song, 2016; Konrad, Trapp, Palmbach, & Blom, 2017; Song, Fisher, Wang, & Cui, 2016), Big Data Analytics even has applications in nearly every field a business manager can think of.

As it is still a research domain, OR utilizes an approach very similar to the way research is conducted in established scientific fields. This process start almost every time by carefully observing and formulating the problem, including gathering all relevant data. This is the first indication that Operations Research and Big Data Analytics are closely related and should be attuned to each other. Next to this, the researcher needs to construct a scientific model that attempts to abstract the essence of the problem. It is implied in and by the model that the conclusions obtained from the model are also valid for the real (business) problem. After the model is developed, suitable experiments can be conducted to test these hypotheses and to validate the model. The tested hypothesis is then withheld or refuted and based on the findings of the experiment, a conclusion can be drawn. *Figure 5* depicts this described OR process (Forgione, 1990).

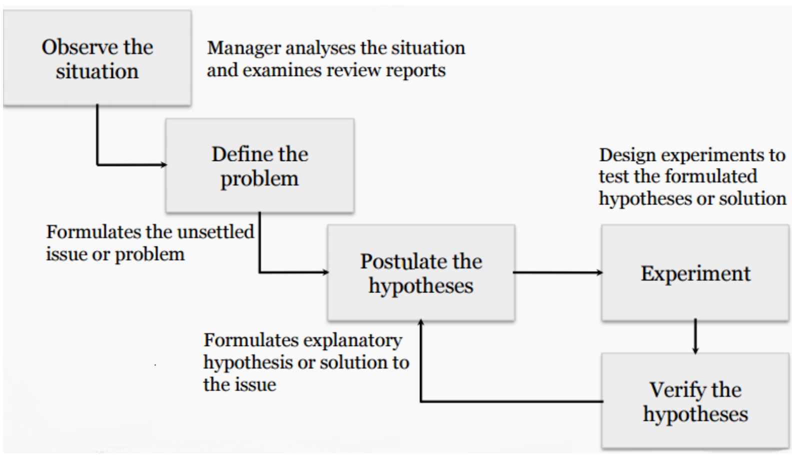

Figure 5: Operations Research Process

OR in general is however more than what is just described. It is namely also concerned with the practical management of the organization. Therefore, OR must provide understandable and interpretable conclusions specifically for the decision makers, who make the final call. Mathematical modeling to analyze complex situations and to make more effective decisions is one of the most popular techniques used in (operations) research. Operations Research is often referred to as Management Science or Decision Science and is then denoted OR/MS. This seems rather trivial, but this extension shows that there is more than research. In the remainder of this dissertation, OR is used to refer to OR/MS. According to the website of INFORMS[2], Management Science is occupied with a number of distinct areas of study including developing and applying models and concepts that may prove useful in helping to illuminate management issues and solve complex problems. It can be executed on **three levels**:

1) A fundamental level that lies in three mathematical disciplines: dynamic systems, optimization and probability theory
2) A modeling level that builds models, gathers data and analyzes them mathematically
3) An application level that has strong aspirations to make a practical impact in the real world

The extension of OR towards OR/MS is important when we go to section 2.3 as the integration of Big Data Analytics in Operations Research makes it more and more a management science where decisions have to be made and insights have to be brought to light.

[2] https://www.informs.org/About-INFORMS/What-is-Operations-Research

2.3 Big Data Analytics in Operations Research

The importance of incorporating analytics into managerial decision making is going up each year. Business analytics has become a great buzz in the operational practice world since Thomas H. Davenport published a series of books and articles from 2006 onwards (Davenport, 2006; Davenport & Harris, 2007; Davenport, Harris, & Morison, 2010). As the amount of data keeps growing exponentially, it is expected that analytics will become more and more important and decisive in the OR approach over time (Choi, Chan, & Yue, 2017; Hillier & Lieberman, 2015). According to Hillier and Lieberman (2015), analytics fully recognizes that we have entered into the era of Big Data, where massive amounts of data now are commonly available to many organizations and businesses to help guide managerial decision making. The primary focus of analytics should be on how to make the most effective use of all the available data, preferably in an efficient way. In a report of the OECD (2013), an annual growth rate between 40 and 60% is even found to be an accurate estimate of the growth of Big Data creation.

2.3.1 Towards a Comprehensive Domain

During the last few years, the disciplines of analytics and Operations Research have been increasingly connected to one another (Brown, et al., 2011). The roots of both areas are quite different, as described in the previous sections, but there are many main elements they have in common. The fact that both domains work with quantitative and most often mathematical models in order to solve real-world business problems is maybe the most visible similarity. People who work in one of both fields regularly have the same background (applied mathematics, industrial engineering, computer science…) or the same interests, so this interwovenness was one of the causes of the growing joint approach. More often than not, the required skills for applicants in both business domains are quite equivalent (Liberatore & Luo, 2010). A grasp of the most recurring skills is: analytical, critical, communicative, problem-solving, math, statistics, computer science… (Marr, 2014). In *section 2.3.2*, we dive into these skills and background of OR and Big Data experts.

If we take a look at the web-based search words, it can be easily seen on *figure 6* that the popularity and searches for Big Data (Analytics) are strongly rising, whereas Operations Research and OR/MS are declining over a period of 13 years (via Google Trends, March 2004 - March 2017). The search peak is depicted on the figures as 100 on the y-axis and every other period in time is measured against this peak, resulting in the following figure:

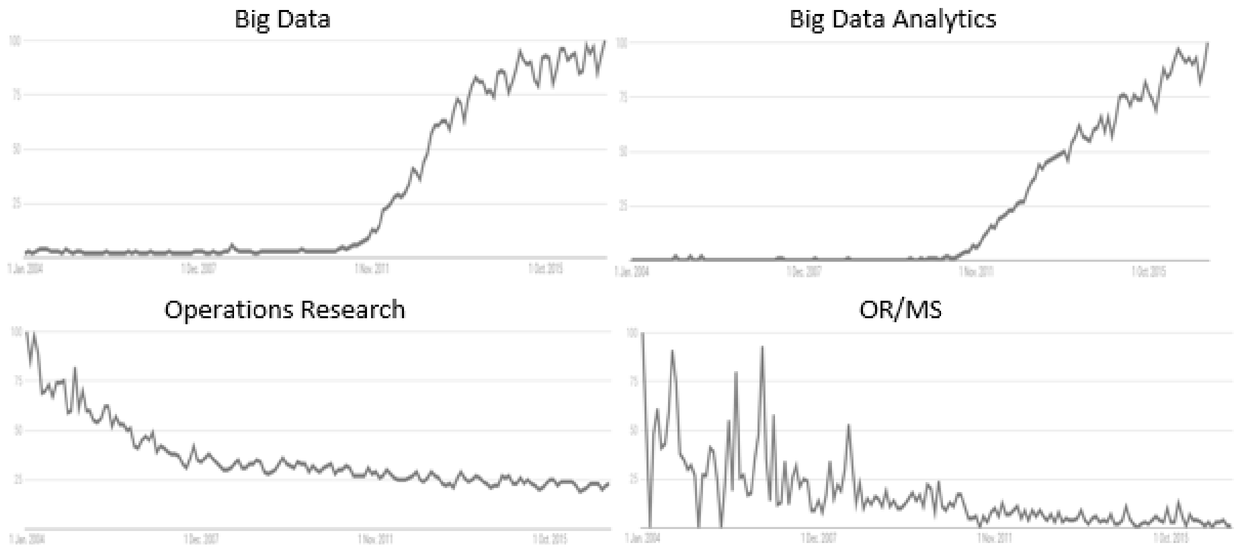

Figure 6: Google Trends Analysis (absolute figures)

What we can conclude from these figures is not that the interest in Operations Research has declined, but that the approach towards Operations Research has changed. A reasonable explanation is that Operations Research and Big Data (Analytics) have become more and more intertwined such that the published and scientific articles about Big Data Analytics relate to the Operations Research business domains. If we look at the relative numbers, we get the following figure (via Google Trends, March 2017):

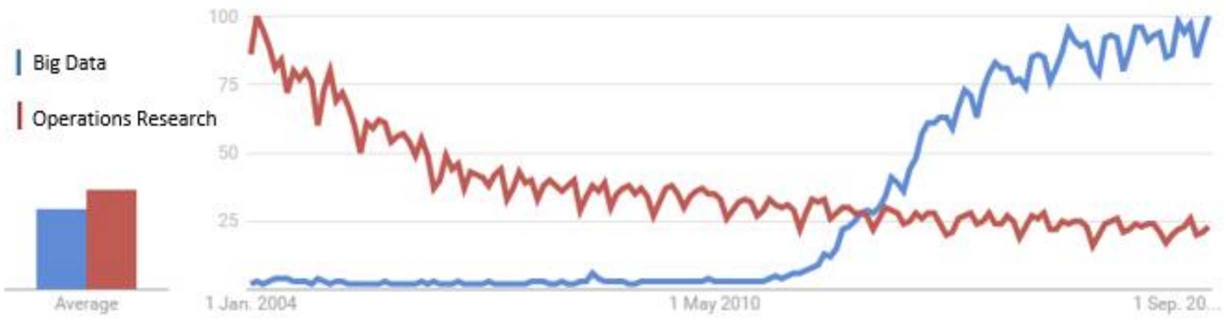

Figure 7: Big Data versus Operations Research - Google Trends (relative figures)

As a field devoted to informed decision-making via advanced modeling (List, 2012), statistical techniques and optimization, it would seem that the practice of Big Data Analytics would fall entirely in the field of Operations Research (Hazen, Skipper, Boone, & Hill, 2016). While OR is not entirely synonymous with (big) data analytics, the work is inherently analytical and to a large extent comparable. According to several authors (Barton & Court, 2012; McAfee & Brynjolfsson, 2012; Waller & Fawcett, 2013), Big Data Analytics and data-driven decisions will even continue to play a big role in OR and that it will be even more powerful than traditional analytics approaches (cfr. data warehouses[3] and spreadsheets). While it may seem a fairly sudden occurrence on the graph, continued innovation of Big Data Analytics in multiple forms has been developing for decades (Acito & Khatri, 2014). The main reason is that data analytics now really has become 'big', changing the nature of how data analytics is performed. Professionals in the fields of Big Data Analytics and Operations Research are slowly becoming aware of one another and their synergy is becoming more manifest (List, 2012). Experts in the two intersecting fields use analogous math modeling and related tools that handle data with advanced software and the flourishing computing power of today's hardware. The aforementioned stages of descriptive, predictive and prescriptive analytics (*section 2.1.2 Analytics*) resonate more and more among Operations Researchers (Haze, Skipper, Boone, & Hill, 2016; Poppelaars, 2011).

2.3.2 Skills

Many studies have investigated the skills employers expect young graduates to have when beginning their career in Operations Research (Sodhi & Son, 2008; Sodhi & Son, 2010). A healthy mix of both technical and soft skills are often in high demand. Employers consistently require modeling, statistics, programming and general analytics skills in an operations management context as their primary requirements. Some less important, though not irrelevant skills, such as communication, leadership, project management, spreadsheet, database, and team skills are often required as well. Many of these skills can be found as requirements for a job in pure data analytics as well. However, to better participate in the Big Data Analytics revolution, OR practitioners need to ameliorate their data management and analysis skills (Liberatore & Luo, 2013). The affiliation with business-oriented skills in process and change management need to be enhanced as well. Required skills for jobs in Big Data Analytics depend on which function the advertisement is trying to fill. Liberatore and Luo (2013) investigated the required skills for three main roles on a large scale: research analysts, application analysts and user analysts. The skills were compared to those required for OR jobs, resulting in the following table (Liberatore & Luo, 2013):

Variable	Analytics		OR		Relative Importance
	Mean	Rank	Mean	Rank	
Data presentation	4.62	1	4.03	5	+ Analytics
Communication	4.44	2	4.14	3	+ Analytics
Business Knowledge	4.44	2	4.08	4	+ Analytics
Problem Recognition	4.41	4	4.62	2	+ OR
Problem Formulation	4.25	5	4.82	1	+ OR

[3] http://searchsqlserver.techtarget.com/definition/data-warehouse

Metrics-KPI Determination	4.17	6	3.69	9	+ Analytics
Interpersonal	4.11	7	3.83	6	+ Analytics
Persuasion	4.02	8	3.80	8	+ Analytics
Project Management	3.96	9	3.83	6	+ Analytics
Change Management	3.71	10	3.44	10	+ Analytics

Table 1: Required Skills Analytics versus Operations Research

We can conclude from this table that the skills required in purely analytics and OR jobs are quite similar, but with a slightly different focus. As mentioned before, this focus will shift more and more towards each other, resulting in more and more congenial job descriptions and skill requirements because jobs will become more of a mix between OR and analytics as well (Mitchell-Guthrie, 2015) - although some authors see analytics as a threat for Operations Research. These skills are of course needed to integrate a (big) data approach into an Operations Research domain. This brings us to the following section, namely the methodology of how to incorporate Big Data in Operations Research.

2.3.3 OR Methodology

In both Operations Research and Big Data Analytics, a unique and systematic methodology is followed to gain knowledge of a certain process or situation. We already sketched both approaches separately in *section 2.1 and 2.2,* but when combined, a whole new methodology finds its origin. Best practices of both approaches are combined in order to create a synergy between (Big) Data Analytics and Operations Research. Their complementarity ensures an integrated approach, a new way of doing business.

According to Chand (n.d) and Rajgopal (2004), OR follows in general the next six steps:

1) Formulating the problem
2) Constructing a model to represent the system under study
3) Deriving solution from the model
4) Testing the model and the derived solution
5) Establishing control over the solution
6) Implementation of the solution

Now, when Big Data Analytics comes into play, this interferes with the above six steps. Once the problem is formulated (*step 1*), a business analysis is performed in order to create an as-is design. This corresponds to a descriptive analysis in the analytics movement and is also closely related to the *second step* of the OR methodology. To derive an optimal solution from the model (*step 3*), huge amounts of data are introduced in the model. Here, Big Data ensures a far more extensive search for the optimal solution. When performing sensitivity analysis (*step 4*), working with a Big Data infrastructure makes sure that the model can be analyzed way faster and more parameters can be investigated simultaneously. After the implementation phase, the model should be monitored continuously. Between every step, there needs to be an open feedback loop, in order to optimally profit from the synergy. *Figure 8* depicts the different steps of this classical OR methodology, from the point of view of Rajgopal (2004). The six steps described above are similar to these.

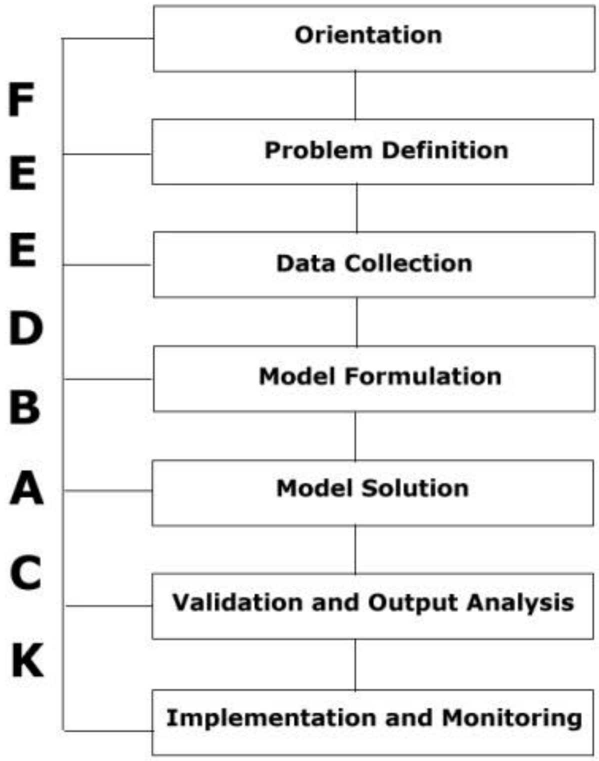

Figure 8: Classical Operations Research Methodology

The classical OR methodology starts with an orientation of the environment and what/where the problem is (at which stage). This results in a problem definition. This definition needs to be very clear, as this is the basis of future research. A wrong or poorly defined problem causes many issues in a further stage and can be quite costly. Based on this problem definition, the right data are collected. Hence, the problem can be implemented in a model that can be objectively analyzed by the data input. Most of the academic emphasis has been on steps 4, 5 and 6, but the reader should bear in mind that the precedent steps are equally important for a practical execution. After the data collection, the model formulation takes place. This is usually the longest phase, as the model is a representation of the system and how it works. There are very different categories of models available, such that each problem can be translated into a model. Once the correct model is implemented, modelers can come to a solution by inputting the data. This solution needs to be validated and a comprehensive analysis of the output needs to be made. This (sensitivity) analysis often says more than a plane solution. The robustness of the model needs to be verified as well. A final phase consists of the implementation of the final recommendation made by the model. Of course, this solution needs to be controlled and monitored.

When walking through the different phases from orientation to implementation, Big Data adds value to the OR model. *Figure 9* depicts this Big Data value chain (CBPL: Commissie voor de bescherming van de persoonlijke levenssfeer, 2016). Data on their own have not much value, as value only arises through the process of collection, preparation and storage, analysis and the usage of these results. Only then will Big Data have economic value. The Big Data value chain is continuously kept in mind while advancing to the next stage in the OR methodology. Hence, the usage of data is optimally integrated into the OR methodology and true value can be created along the Big Data value chain and the OR methodology.

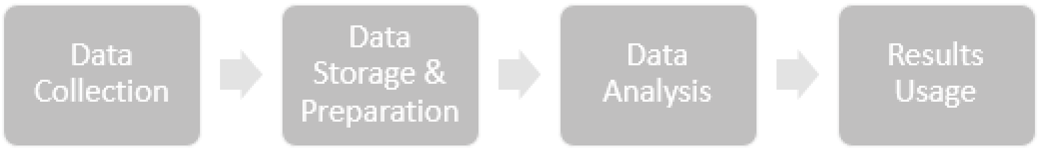

Figure 9: Big Data Value Chain

After the model is implemented and monitored, decisions can be made according to the output and the sensitivity analysis. *Section 2.4* further elaborates on this topic, giving more information of the decision support system. That is the final step before taking concrete actions and trying to gain business advantages. But first, we would like to give a small insight in the existing Big Data Analytics infrastructure.

2.3.4 Infrastructure

Even for the more experienced practitioner of Big Data Analytics, it can be a burden on how to choose the correct tool for his/her projects, let alone for a company on how to choose a complete infrastructure that fully supports the Big Data integration in the daily business of the company (Demchenko, Grosso, & de Laat, 2013). The irony is that the volume and variety of possible technologies is in se inherent to Big Data itself (see *section 2.1.1 Big Data*). *Figure 10* gives an overview of the Big Data Analytics landscape as an illustration of the complexity of choice anno 2016 (Simoudis, 2016).

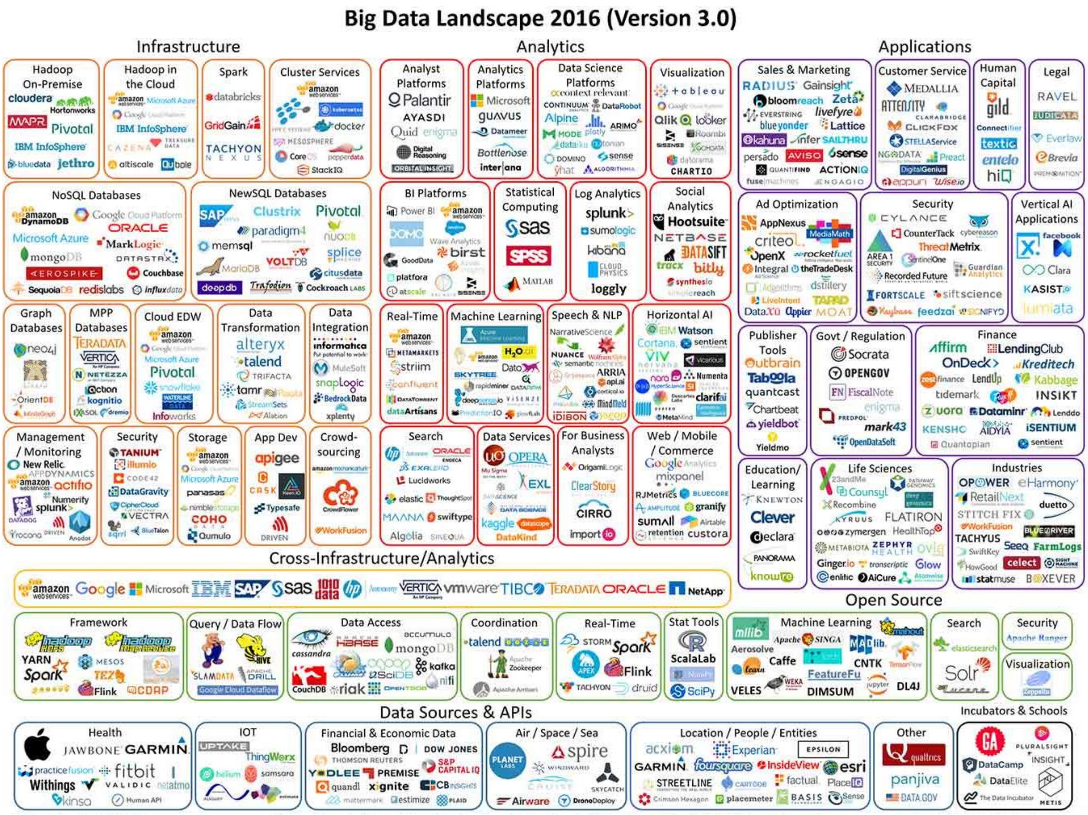

Figure 10: Big Data Technologies

On this figure, you can see there are many possibilities that each support a certain part of Big Data Analytics. In order to integrate many of these applications and software, certain providers have created a platform whereon you can perform these activities using tools. Amazon Web Services[4], Microsoft Azure[5], Google Cloud Platform[6] and IBM Cloud[7] are the biggest cloud providers that enable Big Data (Analytics) tools and software to be used on a large scale by other companies (Panettieri, 2017). More information is available on the adjoined websites.

The aim of this dissertation is to focus on the applications of Big Data Analytics applied in OR domains. As such, we will not dive in further into this tangle of infrastructure possibilities. We refer to Demchenko, Turkmen, and de Laat (2016); Sawant and Shah (2013) and Witten, Frank, Hall, and Pal (2017) for the readers who are interested in exploring the technological options for setting up a Big Data supported infrastructure. One should always keep in mind that data without correct actions are just plain numbers and tokens in a database. Therefore, an adequate decision support system that relies on your data infrastructure needs to be set up. In the following section, we will elaborate on this topic.

---

[4] https://aws.amazon.com/
[5] https://azure.microsoft.com/en-us/
[6] https://cloud.google.com/
[7] https://www.ibm.com/cloud-computing/

2.4 Decision Support System

Before diving in the practical applications and trends of Big Data Analytics in Operations Research (*section 5*), a solid footnote needs to be brought to light: the decision support system. Decision Support Systems (DSS) research has been undertaken for over 35 years and these systems have proven to be useful in supporting unstructured and semi-structured problems (Arnott & Pervan, 2008; Shibl, Lawley, & Debuse, 2013). Today, DSS is often referred to as Business Intelligence (BI), as decision support as a domain has evolved rapidly and picked up the established technological changes.

2.4.1 From data to decision

By combining predictive with prescriptive analytics, people are able to develop new applications in an organization. Taking qualitative decisions is however not solely the result of a data-analysis: an efficient decision support system is indispensable. A decision is supported by (big) data, information, objectives, decision criteria and a decision-supportive model that gives clear insights in all possible alternatives and restrictions of the (business) problem. The role of a DSS is threefold: it predicts a certain aspect of the future with probabilities, it exhibits analytical and modeling capacity and it supports unstructured and semi-structured decisions (Arnott & Pervan, 2014). The continuous process to get to decisions looks like *figure 11* (based on Interaction Design Foundation, 2015). The explanation of the figure is straightforward. But how is this system constructed and which components are essential? The more data input to eventually come to a decision, the more informed the decision. Big Data thus makes sure that there is a larger pool of information available, once there is a context applied to these data. With more information, companies can come to more knowledge or better informed knowledge, if there is a meaning applied to this information. This is where Big Data Analytics comes into play. The extended knowledge can then be transformed into a more extended wisdom, by applying insights to this knowledge. This eventually results in a better decision-making phase when a purpose is applied to this wisdom.

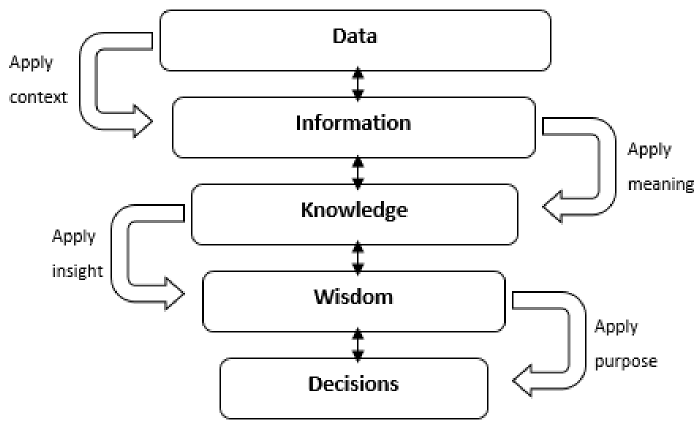

Figure 11: From Data to Decisions - A Continuous Process

2.4.2 Fundamental Components

According to many authors, there are **four fundamental components** of the DSS architecture (Demirkan & Delen, 2013; Haag, Cummings, & Dawkins, 2000; Marakas, 1999; Matthies, Giupponi, & Ostendorf, 2007; Power, 2002). The first and foremost is an easy-to-use user interface. This should be aesthetically pleasing, preferably with a symmetrical layout, appropriate menus and options and easily understandable. The second component is the database. This should serve as the storehouse of information and contains as well internal as external information. Thirdly, a well-built DSS model is required. This model determines how data is analyzed and which information is extracted. Based hereupon, insights are created. The final component is knowledge. This element provides decision makers alternative solutions for a problem and sends signals to managers when there is a significant mismatch between predicted and actual results. These four components are strongly related to the process depicted in the above figure. The main sources on decision support systems date from a while ago (10-15 years). Clearly, decision support systems are not new, but they have come in the digital age as well, and are subject to changes due to the rise of Big Data Analytics (Poleto, de Carvalho, & Costa, 2014). It needs to adapt to the V's of Big Data Analytics, namely volume, variety, velocity, veracity and value (see *section 2.1 Big Data Analytics*).

The DSS should enable a more profound and less time-consuming decision-making process in an organization. An effective DSS provides managers with unbiased data analyses, real-time monitoring and rich reporting (MSG team, n.d.). In general, it makes use of analytical models, several econometric and statistical tools and still depends strongly on human intelligence to be built and to gain insights.

2.4.3 Shortcomings

As always, there are a few issues that can come along when designing and developing an accurate DSS. The most relevant **issues, limitations and disadvantages** are summarized in *table 2*[8]. We can conclude from this table that decision support systems are subject to many issues and limitations. These issues need to be clear when developing a model in order to overcome them. The limitations and disadvantages of the decision support system are mainly due to a lack of integration with Big Data Analytics and Big Data technologies. They are inherent to the limitations of Big Data (analytics) in general (cfr. *Section 6.2 Future Challenges*).

[8] More information can be found on: www.managementstudyguide.com

Issues	Limitations	Disadvantages
Mismatch perspectives programmers – decision makers	Difficulty in quantifying all the data	Information overload
Identification of specific requirements	Unaware of all assumptions	Too much dependence on DSS
Technology selection	Difficulty in collecting all required data	Devaluation of subjectivity
Approach to software design and development	Lack of technology knowledge by users	Overemphasis on decision-making
Fear of learning & implementation		Cost of development

Table 2: Issues, Limitations & Disadvantages of a Decision Support System

The issue of **management acceptance** of the models plays an even greater role in the success or failure of an initiative than the development and implementation of a decision support system. Further literature on this can be found in the bibliography (Levasseur, 2010). This is also a main reason why users do not always take advantage of DSS to support their decision-making (Chan, Song, Sarker, & Plumlee, 2017). The results of the experimental study of Chan, Song, Sarker, and Plumlee (2017) suggest the mediating role of DSS use in the effect of DSS motivation on decision performance. Their results also show that DSS motivation is enhanced in the presence of a more positive DSS performance feedback mechanism, fast DSS response time and high task motivation.

Identifying use cases and user examples related to analyzing large volumes of semi- and unstructured data is still one of the major ongoing challenges for decision support systems and information technology researchers (Power D. J., 2014). Many techniques (such as Hadoop, Spark, Azure…) that are built for the integration of Big Data in OR domains are still developing. It will be a challenge for enterprises to integrate these platforms and software into their daily operations in to capture the benefits of Big Data Analytics in their processes.

The main concepts that are further used in the body of this dissertation should now be clear and well delineated in order to understand the different perspectives that will be used. After setting out the research objectives and methodology in *section 3 and 4*, the applications and trends of Big Data Analytics in Operations Research domains will be discussed in *section 5*.

3. Research Objectives

In the last few decades, Big Data possibilities have been studied widely, in various business domains. There is however no clear literature overview of major trends and applications, categorized by domain. As this is a rather broad challenge, no in-depth analyses will be made. Some business domains will be highly elaborated, whereas others are out of the scope of this paper, because the applications of the four discussed research domains can be found in other business domains as well. Further literature is designated for the interested reader and to maintain integrality.

The objective of this master's dissertation is thus to give an oversight of the most important merging and ongoing trends and applications concerning some important domains of Operations Research that make use of Big Data. The domains that are scrutinized are Marketing (*section 5.1*), Healthcare (*section 5.2*), Operations and Supply Chain Management (*section 5.3*) and Public Services (*section 5.4*). This literature study will give a better insight in Analytics and how this relates to Operations Research. The interaction and interrelationship of these two different domains is key. The implications of these new trends that emerge in OR will be discussed and new research opportunities will be identified as well for Operations Research as Analytics. As this topic is a fast-evolving and continuously developing domain, the objective is to give a snapshot of these trends and applications in a given period of time, as well as a minor sneak-peek into the future opportunities and possibilities that could be further exploited.

Considering the foregoing, the following **research questions** took shape:

1) What prominent trends and applications, according to their research domain, that make use of Big Data are used on a daily basis or are merging in the domain of Operations Research?
2) What are the opportunities for Operations Research as well as Analytics regarding the use of Big Data? Which direction will Big Data take in the future of Operations Research/Management Science?

In line with the first research question (RQ1), we try to give both sides of the story. On the one hand, we will try to demonstrate the benefits of using Big Data and Analytics in Operations Research applications. On the other hand, the limitations and challenges are discussed as well, since every technological revolution always has a downside, so it is self-evident to fill this in in order to captivate the full picture.

The objective of the second research question (RQ2) is to discuss the opportunities and what future lies ahead for the different business domains. How important will Big Data be in the future for companies, governments, citizens…? It is important to sketch the applications landscape the way it is in 2017, but this is not enough. It is far more interesting to also take a look at the future and to predict what is likely to happen. The research on RQ1 can be categorized in descriptive analytics, whereas the research on RQ2 is rather predictive, but well substantiated with the right information and data.

The importance of these research questions will become clear during the reading of this text. There are many sources available but no clear overview exists of several applications, categorized per business domain. Recently, there have been call-ups by some of the most prestigious providers of scientific articles and papers (see *section 6.1 Research Gaps*). On the website of Elsevier, a special issue has been calling for the submission of papers[9]. The aim of this special issue is to collect state-of-the-art research findings on the latest development, up-to-date issues and challenges in the field of Big Data Analytics for business intelligence. Next to the call for papers of Elsevier, there has been a call for papers as well for the 6th IEEE International Congress on Big Data[10]. Research and literature overviews of Big Data, applied in business domains are thus highly wanted, illustrating the importance of these research questions. There are more calls and special issues for the application of Big Data Analytics in real-life businesses, but they can be found in *section 6.1*. The aim of this dissertation is thus to contribute towards a clear and comprehensive overview of Big Data Analytics in OR. The following section will delineate how this will be obtained.

[9] https://www.journals.elsevier.com/expert-systems-with-applications/call-for-papers/special-issue-big-data-analytics-for-business-intelligence

[10] http://www.ieeebigdata.org/2017/cfp.html

4. Methodology

4.1 Big Data Approach

The approach that we followed in this master's dissertation is, by surprise, a **Big Data approach**. All kinds of data, ranging from scientific, published articles to Slidecasts from professors at universities to even blogs from Big Data gurus were collected, screened and analyzed. Here the reader might already observe one of the very unique properties of Big Data, namely the variety of data (cfr. supra). Since the rise of the internet, data has (and this process is still ongoing) grown exponentially. In this large volume of data (the second property of Big Data), we tried to gather the most relevant and scientifically important trends and applications in the OR/MS domain. The approach is not empirical or experimental of nature, but is rather an **explorative and descriptive research**. It is a mixture between a traditional or narrative literature review and a systematic literature review.

Table 3 demonstrates the focus on scientific literature. Next to these papers, websites and documents/reports from websites are the second most important sources. These are especially cited when business applications are discussed, as these are less delineated in scientific literature. Relevant books and book sections are used and cites as well, while business press articles are often used to illustrate the current trends and applications. Several conference proceedings on Big Data, Analytics and Operations Research are consulted and a few highly relevant slides are deliberated.

Type of Source	**Frequency**
Scientific Journal Paper	160
Website	84
Business Magazine Article	38
Book/Book Section	37
Documents/Report	37
Conference Proceedings	22
PowerPoint Slides	5

Table 3: Distribution Type of Sources

Table 4 gives information on the journals used in this dissertation and that are found to be most contributive in literature on Big Data Analytics and Operations Research. These journals are subject-specific, but they often tend to intertwine on subjects related to Big Data and the application of Analytics in OR/MS domains. Other educational important and widely-recognized journals that are used on a single basis are for example International Transactions in Operational Research, MIS Quarterly, Journal of Supply

Chain Management, Journal of Business Logistics, Omega, Journal of decision systems, International Journal of Information Management, International Journal of Market Research…

Journal	**Frequency**
IEEE Internet of Things Journal	8
European Journal of Operational Research	7
Annals of Operations Research	7
Interfaces	6
Journal of Big Data	5
Procedia Computer Science	4
Decision Support Systems	3
Virtual Reality	3
Journal of Automation in Construction	3
Journal of the Operational Research Society	2
European Journal of Information Systems	2
Journal of Cleaner Production	2
International Journal of Advanced Manufacturing Technology	2
Journal of Business Research	2
International Journal of Production Economics	2
International Journal of Operations & Production Management	2
Journal of Marketing	2
Transportation Research Part C	2
Health Affairs	2
IEEE Transactions on Knowledge and Data Engineering	2

Table 4: Journal Perspective (journals with two or more publications)

Next to the technical and scientific literature, managerial literature is assessed as well. This is because the viewpoint of this paper is mainly one of decision making, and this almost always takes place at managerial level. The most important magazines and newspapers in managerial literature used are the following:

Magazine	**Frequency**
Forbes Magazine	13
Harvard Business Review	5
The Guardian	3
Business Insider UK	2
American Banker	2
MIT Sloan Management Review	2

Table 5: Magazine Perspective (magazines with two or more publications)

Focusing on one particular domain is in our opinion not recommended, since it is a key asset of Big Data that it is relevant in any domain one can think of. Therefore, we will discuss the applications and trends in the four following domains: Marketing, Healthcare, Operations and Supply Chain Management (O/SCM) and Public Services. This choice is made on its relevancy as these domains already benefit widely from Big Data integration in the OR approach on a daily basis. We will focus on these four domains, as other OR domains are very comparable to these and conclusions drawn from these four domains can easily reach out to the other OR domains as well.

Since Big Data is continuously evolving and changing rapidly, new sources can be found on a weekly basis. Particular merging trends or opportunities that are described in some sources became outdated by other sources while doing research for this thesis. This has been taken into account as much as possible and we tried to use as many relevant and up-to-date sources as possible, giving a **highly relevant and contemporary overview**. This brings us to the relevancy of the sources used in this literature study (*section 4.2*).

Keywords that were used most often to find relevant sources can be found in *table 6*. In scientific literature, sources were revised on importance via the number of citations and the relevancy score given by the corresponding database/website. The most often used databases are Google Scholar, Web of Knowledge (Web of Science), Elsevier and Crossref.

1.	**Big Data**
2.	**Data Analytics**
3.	**Analytics**
4.	**Operations Research**
5.	**Application**
6.	**Data**
7.	**Challenges**
8.	**Future**
9.	**Operations**
10.	**Decision Support**

Table 6: Top 10 Keywords Used

4.2 Relevancy

We consulted several journals that contribute most to the area of Big Data Analytics usage in Operations Research. Journals such as the European Journal of Operational Research or Interfaces are already given a frequency distribution in terms of relevant articles on this master's dissertation's topic in the above section.

Figure 12 gives an oversight of the sources used per year. It is clear that we focused on the most recent sources, without neglecting the relevant sources before 2013. These figures were finalized in May 2017, therefore 2017 is not proportional to the upward trend after 2013. More recent sources are considered to be more relevant, as the domain of Big Data Analytics is rapidly changing and new applications in OR arise regularly. It has not been quite long since Big Data Analytics is adopted on a larger scale in the industry. Thus, more relevant scientific literature and other sources only started to review and debate Big Data Analytics recently.

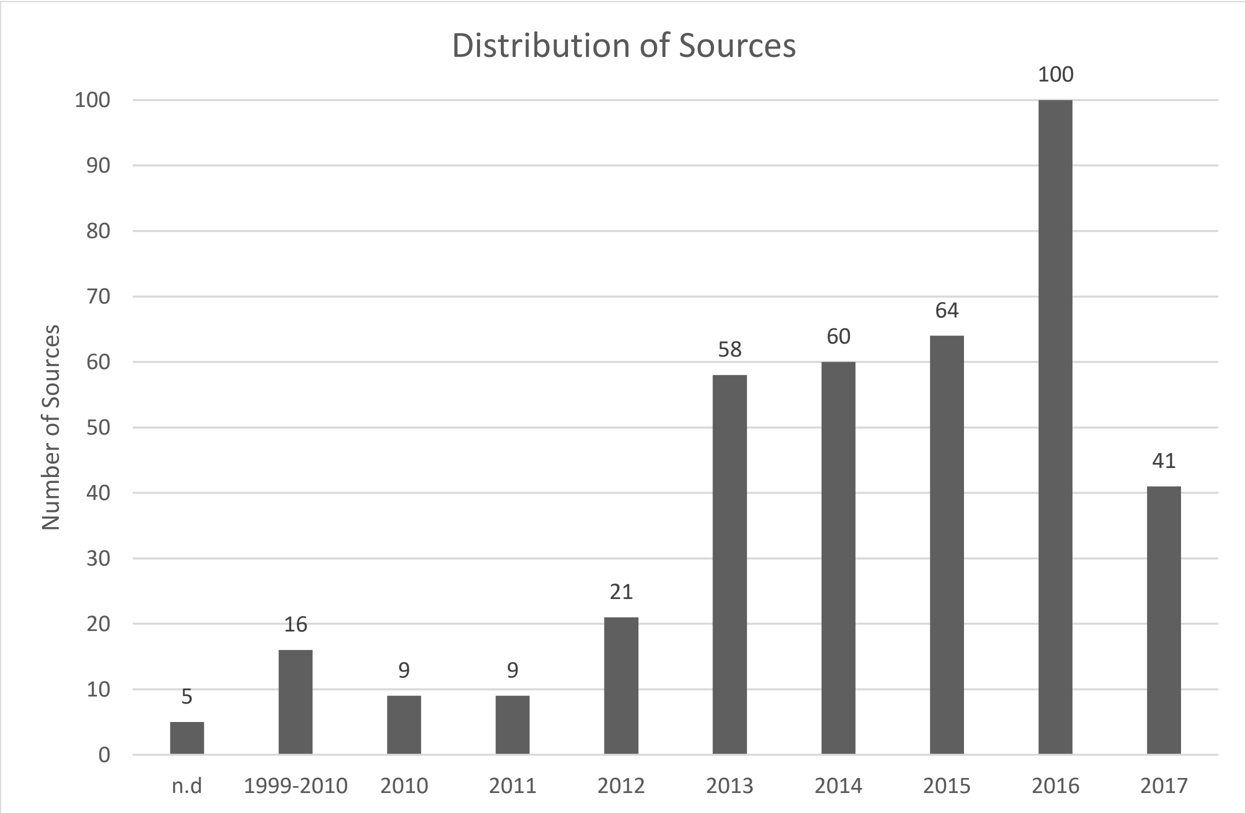

Figure 12: Distribution of Sources

We also present the result of our keyword analysis on the sources used in this dissertation in *table 7*. Such an analysis assists in revealing the intellectual core and identity construction of the discipline under scrutiny, by looking into keywords used in the collected titles of research papers and articles.

Word	**Frequency**	*Word*	**Frequency**
Data	192	*Operations Research / OR*	18
Big Data	160	Business	18
Analytics	64	*Operations*	16
Internet of Things / IoT	33	*Support*	15
Application	32	*Smart*	14
Research	29	*Predict*	13
Challenge	28	*Supply Chain*	13
Healthcare	28	*Augmented Reality*	13
Management	23	*Manufacturing*	11
Decision	23	*Analysis*	11
Artificial Intelligence / AI	23	*Opportunity*	11
Future	21	*Trend*	11
Customer	21	*Decision support*	11
Big Data Analytics	20	*Information*	11
Virtual Reality	19	*Data Science*	10

Table 7: Top 30 Commonly Used Words in Titles (382 sources in total)

It is obvious from the above table that data or Big Data is used in the title in nearly 50% of the collected sources. A reasonable explanation for this is that using Big Data in the title draws the attention of many readers. Analytics is also used very often. Furthermore, we see the various business domains that are discussed in the following chapter, as well as the different perspectives of this dissertation. The large difference between Big Data and Big Data Analytics is probably due to the fact that many authors and writers use Big Data and Big Data Analytics interchangeably. As *Figure 6 and 7* (*section 2.3.1 Towards a Comprehensive Domain*) already indicated, the keyword search for OR and Operations Research is declining. This trend can also be noticed in this table as OR is less used. The keywords that were used most often in the search for relevant sources (*table 5 in section 4.1 Big Data Approach*) are logically covered most in the titles.

5. Applications & Trends

Associated and directly correlated with the emergence of Big Data are the Big Data applications in het Operations Research domains (Chen, Mao, Zhang, & Leung, 2014). According to these (and many other) authors, the amount of data driven applications have grown disproportionately, exceeding Moore's law at the beginning of the new century (Chen & Zhang, 2014). According to Chen and Zhang (2014), excessive data is making great troubles to human beings. But there are so much potential and highly useful values hidden in the huge volume of data, that these troubles will be far surpassed.

A large number of fields and sectors, ranging from business and economic activities to public services, from scientific researches in many areas to national security, have to cope with Big Data problems. A Big Data problem arises due to the inability of conventional database systems and software tools to manage or process the Big Datasets within tolerable time limits (Patel, Birl, & Nai, 2012). Enhanced technologies and a new way of working is required. Therefore, the rise of Big Data applications and technologies also effectuates the rise of new issues. On the one hand, Big Data is extremely valuable to scale up productivity in business and evolutionary breakthroughs in scientific research. On the other hand, many challenges come with the new Big Data technologies. Difficulties in data capture, data storage, data analysis problems and data visualization are but a few challenges Big Data practitioners cope with every day. The immature market of young and talented Big Data students is still in its infancy, there are serious concerns about privacy and ethics of certain data and applications… Thus, we would like to give a clear and comprehensive image of the two sides of Big Data. In each section that is discussed, we first look at the most prominent applications, then we look at the trends that are emerging in the particular domain and at the end we delineate the challenges, issues and problems that occur most often in that sector.

In *section 5.1*, we will discuss which applications are paramount in the domain of **marketing**. Marketing has changed a great deal in the last couple of years (Agrawal, 2016; Fita, 2016), due to the influence of data (Dubois, 2015; Pal, 2015a). Marketers would like to know their customers better and have a complete sketched profile, in order to customize marketing approaches and reach a bigger audience with the right supplies. It is also far less costly to approach people in a personalized matter than to send out bulk advertising. The response rate will be far higher as well when people get advertisements they are interested in, benefiting both parties. The trends that are occurring and emerging are delineated as well.

In *section 5.2*, the **healthcare** domain is discussed as this is one of the largest business areas inside the OR domain. There have been many changes to the healthcare approach, partly influenced by the emergence of Big Data and Analytics (Marr, 2016a). Patient records that were usually stored in a local database are now merged with other patient data that is available (e.g. consumption data, GPS-data…). Like this, a personalized patient care approach becomes possible. Applications in the healthcare area are divided into research, preventive analytics, fraud detection and planning and scheduling. The latter is a widely-discussed field in scientific literature and can benefit enormously from Big Data and Analytics.

We also dive into **Operations and Supply Chain Management (O/SCM)** problems in *section 5.3*. Applications are classified in three large different categories: predictive maintenance, process improvement and risk management. As this domain is also quite extensive, the focus was on the most prominent applications. Data are present at large volumes, in very different formats and are produced very fast, which meets the 3Vs that define Big Data. Machine data, sensor data and process data are combined in order to reduce the risk or cost or to improve output and quality.

Since environmental issues are becoming more and more delicate nowadays (IPCC, 2014), the impact of the **government/public services** on this topic is discussed as well in *section 5.4*. A wide application field are the smart cities, which has been widely discussed in scientific literature but is now gaining foothold in practice. These are cities that try to get the most out of Big Data, Analytics, technology and digitalization to become ‘smart’. A smart city is highly beneficial for the economy, its people, the government, mobility, the environment and the standard of living. The following category of applications are the public data hubs. A hub is a large pool where data is available and whereof public services can make use of to improve its services. A third category is security intelligence. Governments can come to surprising insights concerning security intelligence with the help of Big Data Analytics. Unforeseen patterns can be found, which can improve citizen security.

Each section ends with the discussion of the most important techniques used in Big Data applications in that particular domain. The classification of these techniques is dependent on the nature of the problem and thus the characteristics of the data used in the analysis. The described techniques are not purely used in the corresponding business domain, but are found to be used most often in the most popular and most important applications of Big Data Analytics in that area.

5.1 Big Data in Marketing

When asked to describe marketing in a few keywords, segmentation would almost always make the top three. As customer segmentation is key in marketing, Big Data can help companies to better understand its customers and better divide marketing efforts based on the available data. Traditional marketing efforts already make use to some extent of sales and profit data and rely on this information to anticipate sales and forecast revenue (Paladin, 2016). Knowing that there are approximately 2.5 quintillion bytes of data created each day[11] (that's 2.500.000.000.000 bytes), there is a lot more to capture out of these data than simply using plane sales and profit data from previous years and perform some sort of horse sense logic, as this is an outdated method. The need to extract commercial value and insights from data is not a very new concept, as providing information to help support management insights is considered to be the foundation of the market research sector (Nunan & Di Domenico, 2013). Both established marketing approaches and newer Big Data approaches share the provision of high quality and more timely data into the decision-making process. The key difference however, is that Big Data strategies enable continuous and autonomous decision-making via the use of automation (Yulinsky, 2012).

When asked about their priorities in a Kentico Software research (2015)[12], 54% of digital marketers named Big Data in marketing as their top priority, whereas 56% of marketing professionals saw a revenue increase from data-driven marketing in the first quarter of 2015. Even 80% of senior data and IT decision makers noted a positive change in revenue from using Big Data, showing the relevance of Big Data-driven solutions in the marketing atmosphere.

To introduce the applications section, a **practical example** can already give a bright idea on how to gain value from Big Data in marketing. The 100-year-old direct mail company Harry & David, suffered from the economic recess in the mid-2000s and even filed for bankruptcy in 2011. Today however, they have rising profits, growing customer files and the company is heavily investing in product line expansion. Their secret? They stopped marketing products and started marketing to their customers. Vice-president Paul Lazorisak explained that they rely heavily on data and data-driven solutions and that by using analytics they now know who their customers are, how and when they like to receive offers and who is most likely to increase their spending with the company[13]. Since 2011, their customer retention rate has increased by 14 percent and sales per customer have gone up 7 percent. By segmenting their customers, Harry & David are able to identify which customers are probably going to be most profitable via modeling and

[11] http://www.ibmbigdatahub.com/infographic/four-vs-big-data
[12] http://www.kentico.com/company/press-center/press-releases/2015/millward-brown-survey-shows-strategies-of-successf
[13] https://www.sas.com/en_us/customers/harry-and-david.html

scoring results that stem from both transactional and demographic data sources. Hence, they can focus their attention to their most profitable clients, without neglecting other customers as well. Their efforts are now reasonably divided.

The example of Harry & David is not the only example you can find on companies that can gain valuable insights of Big Data solutions. Besides this approach, there are many other applications of Big Data in the marketing sector, which are described in *section 5.1.1*. In this section, the most paramount applications are delineated and appropriate use cases are given to give a practical interpretation of the application. In *section 5.1.2*, the present and merging trends are discussed and *section 5.1.3* gives an oversight of the issues and main challenges related to Big Data marketing.

5.1.1 Applications

Amongst the many applications of Big Data in marketing, five of the most important ones are discussed in this section. Sentiment analytics, customer 360, customer segmentation, next best offer and channel journey are five of the most prominent applications at this very moment and are discussed in this order. With **sentiment analytics**, firms try to monitor what customers say to increase their marketing success. Organizations also try to identify key customers to boost word-of-mouth marketing and they try to examine customer feedback to improve products and services. **Customer 360** means that enterprises try to identify the customer profile and try to understand the product engagement of that customer. Here, it is also important to detect when a customer is about to leave, in order to enhance the retention rate and make the customer satisfied again, before it is too late (Finsterwalder, 2016). As already explained in the introduction of *section 5.1*, **customer segmentation** is at the top of the agenda for almost all marketers. With a Big Data approach, marketers can design targeted marketing programs, create loyalty programs based on card usage habits and optimize their pricing strategy according each segment's features. Hence, closer relationships can be obtained with valuable customers. As the name suggests, the **next best offer** application enhances customer loyalty through targeted offers (the "best fitted" offers). Like this, increased product propensity is obtained. Another main advantage of this application is that product bundling becomes easier, which uplifts profits. Finally, the **channel journey** provides more relevant content in the preferred channel. The designed model also recognizes multi-channel behaviors that lead (or can lead) to sales. Marketing effectiveness is measured across the different channels during the analysis (Shee, Crompton, Richter, & Maehle, n.d.). After these five applications are discussed, a summary of techniques used within these applications is given.

1) Sentiment Analytics

Since the explosion of social media in the 2nd decade of the 21st century, new opportunities have come to light for organizations to connect with their customers (Scarfi, 2012). The sheer volume of different types of communications across platforms about products, services and brands have been overwhelming companies since a couple of years now. Sentiment analytics can help to quickly read all (or at least most) of this data and even provide a summary of what people like and dislike about a product, a certain service or a company brand. The reasons behind this general sentiment can then easily be extracted, providing managers with valuable business insights after which appropriate actions can be thought out. Text analytics algorithms such as naïve Bayes are ideal for this type of problems. This algorithm analyzes social media posts, documents and feedback folders on its positive or negative sentiment and produces a generic score (Dai & Sun, 2014). Because this data is unstructured, dynamic and ubiquitous, an appropriate Big Data approach is needed. The three main reasons for sentiment analytics are to increase marketing success, to identify key customers and to examine customer feedback. For the latter, the data obtained from the web is combined with internal data such as customer surveys, call logs and customer data in order to get a full and detailed overview of a product's/service's sentiment. *Table 8* gives an overview of the goals, use cases, types of data and sources used in sentiment analysis.

Goal	**Increase Marketing Success**	**Identify Key Customers**	**Examine Customer Feedback**
Use Cases	BBVA, Kia Motors, Nedbank Ltd	Ford, Starbucks, T-Mobile	Barclays, Dell[14], Intuit
Data	Social media, blogs, review sites, customer data, text data, call logs	Social media, demographic data, customer data, purchase data	Social media, feedback forms…
Sources	(Greenwald M. , 2016), (Nedbank Ltd, 2015), (Tena, 2016)	(Erevelles, Fukawa, & Swayne, 2016), (Harris D. , 2014), (Oaks, 2015), (Whitten, 2016)	(Henry, 2016), (Woodie, 2016)

Table 8: Sentiment Analytics Summary

2) Customer 360

If you understand your customer to the highest extent, you can stay ahead of the competition, according to the whitepaper of Shee, Crompton, Richter, and Maehle (n.d.). If a marketer can combine the past and immediate customer behavior, s/he can predict future customer trends and behaviors and what their

[14] http://www.dell.com/en-us/work/learn/business-intelligence-big-data#Related-items

most likely next action will be. When a customer's transactions and travel habits are mapped as well in the same profile sketch, a lifestyle profile can be completed and new insights can be discovered. The first purpose of a customer 360 approach is thus to identify a customer profile. The result is a holistic customer overview and personalized next actions predictions, which is undescribably valuable for companies (Liang Y. H., 2014). A logic consequence of this result are improved marketing campaigns, targeted sales and better customer service because companies better understand product engagement, the second purpose of this approach. In the (hopefully unlikely) event that many customers are on the edge of 'leaving' the company as a customer, the right message or the right compensation can be sent out by the enterprise such that churn rates can stay as low as possible. This is the third main reason for a Big Data approach. *Table 9* summarizes the customer 360 approach.

Goal	**Identify customer profile**	**Understand product engagement**	**Detect customers about to leave**
Use Cases	HDFC Bank[15], HMV[16], OCBC Bank[17], BBVA	First Tennessee Bank, Bank Austria[18], Jeanswest	American Express, Tatra Bank[19], T-Mobile
Data	Customer demographics, transactional data, social media	Customer profile data, transactional data, clickstreams, cookie data	Customer profile data, transactional data, social (media) data
Papers	(Méndez, 2015)	(Cameron, 2015), (Henderson, 2012)	(Coomer, 2016), (Oaks, 2015)

Table 9: Customer 360 Summary

3) Customer Segmentation

Marketers cannot repeat it enough: segmentation is key to efficient marketing efforts. In its essence, segmentation is dividing the customer pool into smaller pools or groups (called segments) that share some similar characteristics or that behave in the same way. Customer segmentation is mainly based on the customer lifetime value (CLTV), a popular marketing concept that has been studied widely during the past decade (Ekinci, Ülengin, Uray, & Ülengin, 2014; Kahreh, Tive, Babania, & Hesan, 2014; Kim, Jung, Suh, & Hwang, 2006). Here, it is essential to build refined strategies for customers based on their lifetime value, as this is profitable for the company. It is thus not a new concept, but in the era of Big Data, it gets another

[15] https://www.sas.com/en_in/customers/hdfc.html
[16] http://is4profit.com/business-blog/big-data-how-closing-the-transactional-loop-could-have-helped-jessops-and-hmv/
[17] https://www-03.ibm.com/software/businesscasestudies/us/en/corp?synkey=J949476V23926U29
[18] https://www.questback.com/uk/case-studies/bank-austria
[19] https://www.sas.com/en_us/customers/tatra-banka.html

dimension. Big Data facilitates to create sharper segments in a more rapid way. It also helps to look at the existing customer base in completely new ways, creating unique business opportunities (Shee, Crompton, Richter, & Maehle, n.d.). The purpose of customer segmentation is fourfold. First of all, it enables to design targeted marketing programs. Every customer segment gets a more or less personalized message thanks to the segmented approach. Secondly, firms can create loyalty programs based on card usage habits. Special offers can be granted to loyal customers by the firm, or by one of the firm's vertical partners. Customer loyalty and per-customer-expenditure can be increased (Brahm, Cheris, & Sherer, 2016). Next to this, firms can optimize their pricing strategy. Every customer has a different financial background and product expectations are distinct as well. The value that customers allocate to a product can be translated in the price they want to pay for it, and this can vary significantly. Big Data can then enhance a sliding scale optimized pricing strategy for the customer base to increase revenues and profits. Last but not least, segmentation via Big Data also enables firms to build closer relationships with more valuable customers. Although not many firms would admit this, they are all guilty of spoiling certain customer segments that are more beneficial for the company, so not every customer is the same in their eyes. Profitable market groups can easily be identified by a Big Data approach and they are given preferential treatment to strengthen the customer satisfaction. *Table 10* provides a summary of the four possibilities of customer segmentation with Big Data, a widely-accepted technique in the marketing world as we know it today.

Goal	**Design targeted marketing programs**	**Create loyalty programs**	**Optimize pricing strategy**	**Build relationships with valuable customers**
Use Cases	Bank of America, First Tennessee Bank, BBVA	Citibank, Royal Bank of Canada, Bank of America	Ryanair, Etihad Airways[20], Fifth Third Bank	Caesars Entertainment, Barclays[21], Zions Bank
Data	Customer data, demographic data	Transactional data, customer profile data	Transactional data, customer profile data, demographic data	Purchase history, demographic data, sales data
Papers	(Groenfeldt, 2013), (Mancini, 2009)	(Groenfeldt, 2013), (Hechtkopf, 2013), (Kirchner, 2012),	(Malighetti, Paleari, & Redondi, 2009), (Passy, 2016)	(Crosman, 2013), (Marr, 2015)

Table 10: Customer Segmentation Summary

[20] https://datafloq.com/read/etihad-airways-big-data-reach-destination/412
[21] https://www.barclayscorporate.com/content/dam/corppublic/corporate/Documents/insight/Big-Data-report.pdf

4) Next Best Offer

An interesting question for companies is how they can upscale their sales. On the one hand, a company can try to increase its sales by reaching out to more customers. On the other hand, the company can try to up-sell and cross-sell its products with their existing customer base. The latter can be achieved by predicting what the customer exactly wants. This is possible thanks to a Big Data approach called '*the next best offer*'. In this approach, customer's market baskets are analyzed and marketers try to find patterns between products to forecast future purchases for all customers individually, but en masse. This knowledge can then be leveraged by decision makers to improve the return on investment by marketing efforts, but higher customer loyalty and increased sales results can also be a direct outcome. The best uses for '*next best offer*' are threefold. First of all, it enhances customer loyalty by offering the customer what s/he (probably) wants. Relevant offers might increase the interest and customer stickiness to the brand or company (Shee, Crompton, Richter, & Maehle, n.d.). Secondly, the purpose of next best offer can be to measure product propensity. This knowledge is also used to boost revenue via personalized product/service offers. Finally, it also enables product bundling to uplift revenue. Once determined which products are most likely to be purchased together, it becomes a matter of good execution and targeted promotions to increase revenues. Big Data is also used to increase customer involvement in the product development stage, as customer needs are incorporated. This also enhances customer loyalty and facilitates the bundling of products to uplift revenues (Zhan, Tan, Li, & Tse, 2016). *Table 11* provides a concise summary of the purposes of the '*next best offer*' application, along with some interesting use cases.

Goal	**Enhance customer loyalty**	**Measure product propensity**	**Bundle products to uplift revenues**
Use Cases	Tesco, Netflix	Westpac, Amazon	Eircom
Data	Transactional data, customer profile data	Purchase history, customer profile data, transactional data	Transactional data, customer profile data, product data
Papers	(Marr, 2016b), (Nelson R. , 2015)	(Beuder, 2013), (Corner, 2014)	(Figueras & Mayer, 2012)

Table 11: Next Best Offer Summary

5) <u>Channel Journey</u>

Another buzzword in marketing is the 'customer journey'. This relates to the stadia customer go through in their relationship with a certain brand or company (McColl-Kennedy, et al., 2015). In this era of increasingly complex customer behavior, it is key for companies to maintain a close relationship with its customer throughout the whole journey, to accompany the customer towards a better outcome for both him/her and the company, creating a synergy. The increasing focus on customer experience throughout his/her journey arises because customers now interact with firms through myriad touch points in multiple channels and media, resulting in more complex customer journeys called channel journeys (Lemon & Verhoef, 2016). Keeping track of the customer journey can be a very difficult task, as customers can interact with a company via mobile, social media, brick-and-mortar stores, click ads, publication platforms, television… Big Data can give a helping hand by taking a holistic view of the entire customer journey and corresponding experiences on each channel. This can identify patterns of usage that eventually lead to sales, at what point of the customer journey it might possibly go wrong or which channels are underperforming (Shee, Crompton, Richter, & Maehle, n.d.). The purpose of a Big Data approach along the customer journey is fourfold. First of all, it enables firms to provide more relevant content in the preferred channel. Not all channels are used by the same audience or with the same intentions, so the message a firm sends out on these channels should be adapted accordingly. Secondly, this approach facilitates the recognition of multi-channel behavior that (eventually) leads to sales. Optimization of funnel conversion becomes possible once these patterns are exposed. Next to this, a Big Data approach permits a company to guide customers to low-cost channels. Not every channel comes at the same cost for a company, while they serve the same purpose. This requires however a robust self-service and a customer-driven approach to provide a qualitative and intuitive service on the low-cost channel (Shee, Crompton, Richter, & Maehle, n.d.). Finally, marketing effectiveness can be measured across all existing channels. It can be hard to determine which channel can take credits for 'making the sale'. It is perfectly possible that a customer sees an advertisement at a certain channel and incited by this advertisement, but s/he can buy a product at another channel. Therefore, marketing effectiveness is measured in its entirety. *Table 12* provides a brief summary of using a Big Data approach during the channel journey of customers and which purposes this serves.

Goal	Provide Relevant Content in Preferred Channel	Recognize Multi-Channel Behavior	Guide Customers to Low-Cost Channels	Measure Marketing Effectiveness across all Channels
Use Cases	HDFC Bank, OCBC Bank, Bank of China	Vodafone, GE, MoneySupermarket	HSBC, UniCredit	Bank Polski[22], Laurentian Bank of Canada
Data	Cookies, clickstream data, URL referrals	Customer profile data, cookies, tracking codes, online JavaScript data	Online JavaScript data, time stamps, customer profile data	Clickstream data, tracking codes, cookies, customer profile data
Papers	(Guerrieri, 2014), (Ramshaw, 2011), (Vasudevan, 2016)	(Experian, 2013), (Wainewright, 2016), (Winig, 2016)	(Kumar A. , 2015), (Splunk, 2016)	(Laurentian Bank, 2016), (Story, O'Malley, & Hart, 2011)

Table 12: Customer Journey Summary

6) Techniques

The five prominent applications discussed above are all disruptive in the marketing area. During the past decade, they have proven to be very lucrative when implemented and handled correctly. The influence of decision makers stays important, but now they are supported by the model/system, which gives insights, patterns and recommendations. To develop this system, some complex algorithms are required. There are **six important techniques** that companies make use of most often: decision trees & random forests, clustering, text analytics, neural networks, link analysis and survival analysis. What follows is a brief introduction to the particular technique and why it is an important technique in many marketing applications.

- **Decision trees** are one of the most powerful data mining techniques, as they can handle a diverse array of problems (Shee, Crompton, Richter, & Maehle, n.d.). Almost any data type (nominal, numeric or other types) can be handled by decision trees easily, which is ideal for Big Datasets. Typically, a decision tree classifies data into a predefined target field by splitting up data into smaller data cells. **Random forests** boost the efficiency of decision trees by creating different trees that model the same target, all being slightly different. This diminishes the possible errors and noise of an individual tree model. *Figure 13* gives a classic illustration of a decision tree model

[22] http://www.teradata.com/News-Releases/2012/PKO-Bank-Polski-and-Teradata-Deploy-Multi-Channel-Campaign-Management-Platform/?LangType=1033&LangSelect=true

for a small dataset (Marchi, 2013). The flow of the data is easily recognizable and hence, Big Datasets can be classified into smaller cells. The result are small data cells containing the right data. This makes it easier to analyze the problem.

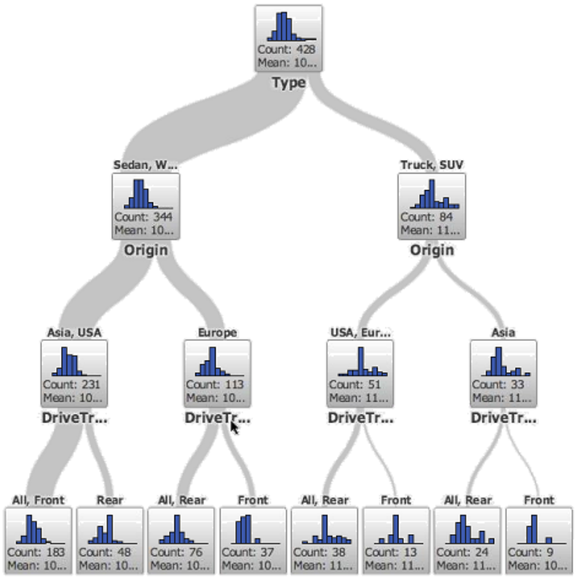

Figure 13: Decision Tree

- **Clustering** in this context basically means the automation of finding meaningful patterns within a data set. A common problem in data mining is identifying which patterns are actually useful in the pool of all discovered patterns in the data. In the marketing context, clusters are nearly always called segments, so this is one of the most commonly used applications of the clustering technique as segmentation is key in marketing. A popular example of a clustering algorithm is the K-means algorithm (Celebi, Kingravi, & Vela, 2013).
- **Text analytics** has been widely accepted in organizations as there has been an explosion of generated text both internally and externally along with the rise of social media and online services (Singh, 2016). This technique automates the reading process and provides a brief summary of thousands of documents at a fast pace. It relies heavily on probability theory and the uncommonness and occurrence of certain words, as this determines the meaning and the themes of the written texts (Hu & Liu, 2012).
- **Neural networks** are based on the networks found in biological studies on the human brain. Here, nodes are activated by a signal that in turn transmits a response signal to activate other nodes, thus forming a network of nodes or a neural network (Shee, Crompton, Richter, & Maehle, n.d.). The total received signal is calculated per node, based on certain appointed weights. The target field of the algorithm is typically a scoring function (for example the customer propensity towards

a particular product or service). The outcome of this technique is that the most likely action of a customer can be determined, given a certain action/event triggered by the company.

- Similar to the neural networks is the **link analysis technique**. As the name suggests, this technique searches for relationships and connections (= links). Originally, this technique is part of a subset of mathematics called graph theory. It digs into the existing data and objects to find links and patterns (Berry & Linoff, 2011). A successful application of the technique is in identifying key sources of information on the internet by analyzing links between pages, comparable to following clickstream by cookies.
- **Survival analysis** can signal a company when to start worrying about an event. This is often applied in healthcare (patient survival rates) and manufacturing (failure rates), but is climbing its way up to marketing applications. It is very important for marketers to know more or less when a customer is going to leave the service of a company or when s/he is going to stop buying a certain product. In combination with purchase history and based on these findings, the customer lifetime value can easily be determined and companies can accordingly undertake certain actions or let things go the way they are (Moore, 2016).

5.1.2 Trends

General

A major shift, turning the scientific method around from fitting data to preconceived theories of the marketplace to **using data to frame theories**, has been occurring since 2011 (Firestein, 2012). It is comparable with shifting from a more deductive method towards an inductive approach of scientific research. Besides this, methodological and technological advances enable researchers to discover patterns in Big Data without formulating hypotheses (Lycett, 2013). Hence, it becomes easier for researcher to focus on the data and explore the unknown than to focus on what is already known. As Erevelles, Fukawa, & Swayne (2016) see it, quite some time can be saved by not always confirming what has already been confirmed and more new findings may arise, perhaps slightly changing the existing theory or the established business approach.

Another major ongoing trend in the rise of Big Data that not only affects the marketing area, is the **Cloud** emerging as a preferred deployment model for Big Data (Pentaho, 2016). Big Data applications will move to the cloud to reach out to even higher volumes of data at a speed that is never seen before. The pay-per-use charging of the clouds offered by several business giants lures companies towards the cloud more and more. Companies are only charged for the space and applications they use inside the cloud, as well as how long they store their data. There is thus not an exuberant initial investment in Big Data infrastructure, as everything happens online on distributed computer networks from the cloud provider, brought to other companies via their cloud system. The issue of privacy becomes a very urgent challenge and is expected to be the biggest challenge of Big Data in the future (Gahi, Guennoun, & Mouftah, 2016; Tene, 2013). *Section 6.2.1 Privacy* delineates this (future) challenge.

Specific

As the speed of technology, competition, inputs and the markets in general increases, the speed and complexity at which a firm acquires and analyzes information must also increase (Xu, Frankwick, & Ramirez, 2016). The research of Xu, Frankwick, & Ramirez (2016) suggests that all firms will benefit from using and combining both Traditional Marketing Analytics (TMA) as Big Data Analytics (BDA). However, not all situations and circumstances justify the cost of collecting and analyzing both types of data or having two separate technologies on point. Therefore, with the rise of Big Data technologies and infrastructure, it is recommended in their paper that a comprehensive and circumstantial analysis is performed in order to determine the correct data approach.

There are four common Big Data trends in marketing that recur in many articles:

- First and foremost, **Big Data talent shortages** become more and more critical. Where Big Data continues to grow and becomes more and more mainstream in the marketing domain, there is a growing shortage of skilled young talents to fill in the Big Data jobs, creating scarcity on the market. Universities, colleges and other institutions are desperately lagging behind, trying to fill in this gap by offering more and more data science programs, but right now there is an intermediate stage with massive Big Data investments and a severe shortage of adequate personnel (Davenport & Patil, 2012).
- A second major trend is that **analytical tools become more mature and integrated** in organizations. There are many Big Data tools and applications out there for marketing purposes, but they often still lack a certain degree of integration. The focus on easy integration and improved maturity can easily be seen when looking at which new tools come out each year. Data analytics is not a new concept anymore and the tools and applications become more and more ready for massive adoption, ready to create value for more organizations. The hype has now settled down to find its way to success in the real business world (Someh, Wixom, Davern, & Shanks, 2017).
- A third important trend that is noticed by several authors, is that Big Data marketing tools have been of a predictive nature until now. It is predicted (pun intended) that these tools will go a step

further, towards **prescriptive analytics** (see *section 2.1.2 Analytics*). They'll advice the marketer what to do next, not only what's going on (Wedel & Kannan, 2016).

- The fourth major trend brings us directly to the next section of issues and main challenges. Big Data marketing faces **increased challenges related to consumer privacy**. Many customers have been more or less accepting the intrusiveness of marketers up until a certain extent, but if they push it too far, a public recoil is peeking around the corner (ReachForce, 2016). Therefore, marketers should handle their data carefully and in a respectful way because a public collision can have disastrous consequences in the digital era.

5.1.3 Issues & Main Challenges

As mentioned in the previous paragraph, a common issue in Big Data is the gray zone between ethically correct and ethically incorrect. More often than not, it is not quite clear where the company stands in the gray zone at a certain moment in time. **Privacy** is a subjective feeling and varies widely from person to person. Some people don't care that marketers keep track of purchase history and make customer profiles based on social media and online behavior. However, many people are not eager to give up their privacy for marketing purposes, although they have open social media profiles whereof marketers can (under certain conditions) subtract information. Privacy rules and regulations vary from country to country, from medium to medium and from platform to platform, but also from company to company. In this tangle of rules, companies tend to be very creative in following these rules, interpreting them to their own best interest. In *section 6.2.1 of Future Challenges*, we will go into detail of the privacy issue, with a clear difference between European and American regulations.

It is particularly hard for a company to **gather relevant information**. Many data they dispose of are irrelevant and can even interrelate with other data, resulting in found patterns that are not very useful (Chen, Mao, & Liu, 2014). It is one of the main challenges organizations face today, because there is an overload of data available. To keep the dataset rich and valuable, databases and data lakes need to be cleansed while retrieving new data online.

The advantage of the many tools out there can also become a challenge for an organization, as it becomes not quite clear **which analytical tool is best suited** for the organization to fulfill a certain purpose. There has even been a Big Data-based research on how to select the right tool (Wang, Zhao, Liu, & Chen, 2017). Companies however agree that they are better off when they are able to choose between more alternatives than they have a modest choice.

The most important challenge for Big Data usage in marketing is knowing **how to go from data to insight to impact**. Regardless how much data you have or how well you analyze them, without taking the right decision it will not have a very big impact. Nevertheless, a huge impact is the main reason for companies to invest in Big Data adoption (Farah, 2013). Whether you create a higher customer experience, more sales, closer customer relationships or whatever impact it may be, this impact can only be achieved when taking the right actions. It is thus immensely important to interpret the Big Data analysis well and have a common sense about doing business right. Only then will Big Data Analytics pay off. A marketer should keep three things in mind when using Big Data: s/he should use Big Data to dig for deeper insights, get the insights to the right people and keep focused on a few key objectives instead of doing all at once[23].

[23] https://www.sas.com/en_us/insights/big-data/big-data-marketing.html

5.2 Big Data in Healthcare

It is a known fact that people nowadays live longer and in better circumstances. But recent technological developments will enable us to improve our health even more (Philips, 2016). If you are asked to think about some small steps that have been taken in healthcare, you immediately can think of devices such as step counters, wearables or electronic patient files. But these are just the beginning of a disruptive wave that is flooding over the healthcare industry. These examples mainly focus on measuring small things that can be linked to a patient's lifestyle or his/her physical abilities. Although this may seem unimportant or trivial, the so-called 'quantified health' is going to be of vital importance according to many doctors (Wang, Kung, & Byrd, 2016). They will be able to collect targeted health records from patients, which facilitates the process of prevention. Preventive healthcare will become the standard of the future (Batarseh & Latif, 2015), and we all know prevention is better than to cure. With the tremendous increase in adoption of Electronic Health Records (EHR), various sources of information become available about patients across hospitals. The challenge in using these EHRs is in identifying the appropriate and effective uses of EHR data to improve patient safety in early stages without requiring additional effort from physicians, so they can focus on their work (Sun, Wang, Hu, & Edabollahi, 2012).

Another main advantage of Big Data in healthcare is that millions of patient records can be kept up in a large database and an analysis on these historical data can give insights in overlapping symptoms, common patterns and certain disease trends that may have been never even considered before. The KU Leuven for example is keeping track of anonymized patient data in a Big Data environment in order to come to new insights in different illnesses (Van Nieuwenhove, 2012). Lab data, prescribed medication data and diagnoses of general practitioners are the source of finding epidemiological trends and prescription behavior of doctors is mapped and analyzed. If these data are connected to medical files at hospitals in a cloud system, the advantages are unforeseen. A common discussed downside of this approach is that if this cloud system is hacked, violently abused or there is a leak, many companies can benefit from these data for profitable purposes[24].

A nice example of how Big Data could impact public health can be found in the paper of Khoury and Ioannidis (2014). In 1854, the father of modern epidemiology doctor John Snow recorded the location of affected houses during the cholera outbreak in London. After quite some time and hard work, he implicated the water pump at Broad Street to be the source of the outbreak, without even knowing what the cholera organism was. With today's technology, Snow might have utilized GPS info and disease prevalence data to solve the problem within hours. That is how diseases are mapped nowadays, to

[24] https://www.helpnetsecurity.com/2017/02/23/healthcare-data-breaches/

prevent further collateral damage (Drexler, 2014). Cooperation with experts in the fields of biology and bioinformatics is also very admissible and thus it is unlikely that such a large-scale epidemic could take place in a metropolis in 2017.

This example demonstrates the importance of public health data in order to prevent large-scale diseases. On a smaller scale however, it is also of upmost importance to keep track of health data for every person individually in order to improve living standards in terms of personal healthcare. There are many applications that make use of Big Data techniques in healthcare, all serving a different purpose and with a different angle of incidence. In *section 5.2.1*, the most paramount applications are delineated and appropriate use cases are given to give a practical interpretation of the application. In *section 5.2.2*, the present and merging trends are discussed and *section 5.2.3* gives an oversight of the issues and main challenges related to Big Data in the healthcare environment.

5.2.1 Applications

A huge variety of health data is being collected at a very high pace (Murdoch & Detsky, 2013). In the healthcare sector, Big Data analyses are used for very distinct purposes. According to Ottenheijm (2015), the applications can roughly be divided into two categories: business intelligence and medical intelligence. Applications that handle data for daily operations fall under the first category, as these are the business analytics applications. An example of this category are digital dashboards with insightful presented data for managers and directors of the hospital. The latter category, medical intelligence, includes data that are used for the purpose of patient care and scientific medical research. Decision supportive software or predictive analytics for patient treatments are examples of medical intelligence applications. In the remainder of this section, **four application areas** will be discussed, along with some practical use cases or examples to draw a clear picture of the Big Data healthcare landscape. First of all, the **research opportunities** are discussed. Research is essential in healthcare as behavior and lifestyle changes over time, so do diseases, infections and symptoms. Findings in the healthcare industry are universal and can save thousands of lives, so it is of the upmost importance to conduct research in the best way possible, preferably at a high pace. Secondly, **preventive analytics** is discussed. The saying goes ‘prevention is better than cure’ and this is certainly the case in hospitals. Patient safety is on the top of the agenda of (almost) all hospitals. If there is a slight indication of a certain disease due to specific symptoms in combination with the patient’s history and lifestyle, worse can be prevented with analytics. It is always better for doctors to get to the patient (or the other way around) in an early stage. Next to this, **fraud detection** is a main application of Big Data in healthcare. Fraud still counts for a significant amount of money in the healthcare industry and Big Data technologies can put an end to this by evincing irregularities and fraudulent transactions. Finally, **planning and scheduling** gets a whole new dimension in the Big Data era. More data is used in scheduling personnel, based on expected patient admission rates and historical data. The likelihood of returning patients is considered, as well as popular illness periods. Hence, real-time optimization of nurse/doctor/operation room scheduling becomes the standard with a Big Data approach. After these application areas are discussed, two important techniques used within these applications are delineated. This is to concretize the way of operating of these applications.

1) Research Opportunities

The use of Big Data for healthcare discoveries in science is disruptive in comparison with traditional scientific research. Traditional research first formulates a research question and then collects data according to certain predefined samples to test the hypothesis. The usually modest amount of data is then analyzed to ascertain a causal relationship between two characteristics. The reverse is true for Big Data. With this new approach, data is collected first in considerably larger amounts with the aim of searching for expected and unexpected connections, relationships of links. Because of this reversed approach, there are a lot more findings than a sampled-approach could ever find. Research towards the functioning and impact of medicines follows this Big Data approach (Krumholz, 2014). Public health is thus largely influenced by these renewed research opportunities, and this is the main purpose of using Big Data in healthcare environments. *Table 13* summarizes the goals with appropriate use cases and types of data.

Goal	**Research**	**Public Health**
Use Cases	BRAIN[25], Head Health Challenge	AEGLE, Yoda project, UK Biobank, PatientsLikeMe
Data	Social media, medical records, unstructured clinical data, population behavior data, genomic data, structured EHR	Social media, medical records, historical data, patient data, population data, structured HER, genetic data
Papers	(Oztekin, 2017), (Sun & Reddy, 2013)	(Gesundheit Österreich Forschungs- und Plannungs GmbH, 2016), (Sun, Wang, Hu, & Edabollahi, 2012)

Table 13: Research Opportunities Summary

[25] https://obamawhitehouse.archives.gov/BRAIN

2) Preventive Analytics

With the introduction and further development of smartphones, a new world has opened for collecting data on behavior, lifestyle and health (Ottenheijm, 2015). Handy applications change mobile phones into step counters, sleep monitors or even medication guards. Another development is the rise of wearables (think of Fitbit, Withings, Jawbone or Samsung Gear Fit) that users can utilize to measure, save and compare their behavior and certain health values with others. Users can even share this data with their doctors, enabling the latter to get a better image of the patient. Hence, personalized patient care becomes possible. By combining the information of treatment outcomes with the patient's characteristics and health values, treatments can become more and more personalized. Like this, diseases can be treated far more effectively than before (Chawla & Davis, 2013). Another purpose of healthcare analytics is to prevent the outbreak of large-scale epidemics, called epidemic control. While tracking the spreading and distribution of infections and analyzing illness patterns, public health can be improved by an increased reaction speed to prevent further outbreak. This facilitates the deployment of assistance in the affected areas in an effective and efficient way. *Table 14* gives a brief overview of available use cases for each purpose of preventive analytics.

Goal	**Personalized Patient Care**	**Epidemic Control**
Use Cases	Propeller Health, Project Artemis, Center for Personalized Cancer Treatment (CPCT)	Epidemic in Vellore (India), Google Flu,
Data	Patient health values, medical records, public health data	Social media, public records, population data, medical records
Papers	(Katz, 2016), (Lonzer, 2015), (Proffitt, 2012),	(Lazer, Kennedy, King, & Vespignani, 2014), (Lopez, Gunasekaran, Kaur, & Abbas, 2014)

Table 14: Preventive Analytics Summary

3) Fraud Detection

The presence of Big Data techniques allows organizations to easily detect financial irregularities and fraudulent practices. An investigation of the Dutch Care Authority showed that medical specialists, general health practitioners, nursing institutions and other caregivers had filed remarkable declarations for 800 million euros in 2012 (Nederlandse Zorgautoriteit, 2014). They invoiced significantly more hours than colleagues or let treatments run a little longer until they got a higher compensation. Estimates of the financial damage in the healthcare in the Netherlands diverge from tens of millions towards billions of euros each year (de Bruijn, 2013). The National Health Care Anti-Fraud Association in the USA even estimates that three percent of all health care spending, corresponding to 60 billion dollars and more, is lost to healthcare fraud (Isbitts, n.d.). When detecting fraudulent practices, certain electronic mistakes can come to light. These can be caused by human mistakes and such, quality control gets a new dimension. The fraud detection software makes sure that there is an extra mechanism that checks whether everything is filled in right in the electronic data files. *Table 15* summarizes the main applications for fraud detection.

Goal	**Fraud Detection**	**Quality Control**
Use Cases	Center for Medicare and Medicaid Services (CMS), National Health Care Anti-Fraud Association	UnitedHealthcare, the Transformed-Medicaid Statistical Information System (T-MSIS)
Data	Electronic health records, accounting data, patient data, workflow data	Electronic health records, accounting data, patient data, workflow data
Papers	(Brennan, Oelschlaeger, Cox, & Tavenner, 2014), (Isbitts, n.d.)	(Barclay, n.d.), (Menon & Sheth, 2016)

Table 15: Fraud Detection Summary

4) Planning & Scheduling

The complexity of healthcare systems is often terribly intricated. Personnel schedules and rostering, patient scheduling, room scheduling… In an ideal hospital, these are all part of the same integrated system. The efficient and effective management of the personnel is of critical importance in a healthcare environment, as this counts for a vast share of the operational costs (Maenhout & Vanhoucke, 2013). A Big Data approach may however reduce costs in a hospital by leveraging its analytics. According to Harris, May, and Vargas (2016), accurate predictions of no-shows by patients may assist a clinic in developing operational mitigation strategies (such as overbooking appointment slots and special patient management). They developed a new predictive model to produce probability estimates. Based on this analysis, operational scheduling can be adjusted accordingly. Managing the flow of patients through a Big Data approach is far more optimal than the way it is conducted today. Hospitals are on the right way, but mainly use historic data. However, we have to look at the future and predict what will happen then, not solely based on how things went earlier. Patient experience is also hugely impacted by the Electronic Health Records (EHRs) that are available nowadays, enabling doctors to come to a diagnosis more rapidly thanks to the quick and easy access to patient data. *Table 16* is a concise summary of important applications in planning and scheduling.

Goal	**Optimize Personnel Scheduling**	**Improve Patient Experience**	**Cost Reduction**
Use Cases	Assistance Publique-Hôpitaux de Paris (AP-HP), Attendance on Demand	Cleveland Clinic, Apollo Hospitals	Heritage Health Prize
Data	Scheduling data, patient data, historical data	Medical data, patient data,	Hospital admission data, patient data, medical data
Papers	(Attendance, 2015), (Marr, 2016c)	(Jha, 2016), (Pal, 2015b)	(Sun & Reddy, 2013)

Table 16: Planning & Scheduling Summary

5) Techniques

The four prominent application areas discussed above are all disruptive in the healthcare area. During the past decade, they have proven to save lives, improve public health and prevent further epidemic spread in affected countries. There are two very important techniques that hospitals/researchers make use of most often: machine learning and text mining. What follows is a brief introduction to the technique and why it is an important technique in many healthcare applications.

- **Machine Learning** is a method of data analysis that automates analytical model building. It is the science of getting computers to act without being explicitly programmed[26]. It improves diagnostics, predicts better outcomes and is revolutionizing personalized care. A great example is the IBM Watson that saved the life of a woman dying from cancer (Bort, 2016). The Watson computer system ran her genomic sequence and found it she had two strains of leukemia instead of the discovered one. This enabled another and more substantiated cure. The process of continuous learning for machine learning is depicted in *figure 14* (Kearn, 2016).

[26] https://www.coursera.org/learn/machine-learning

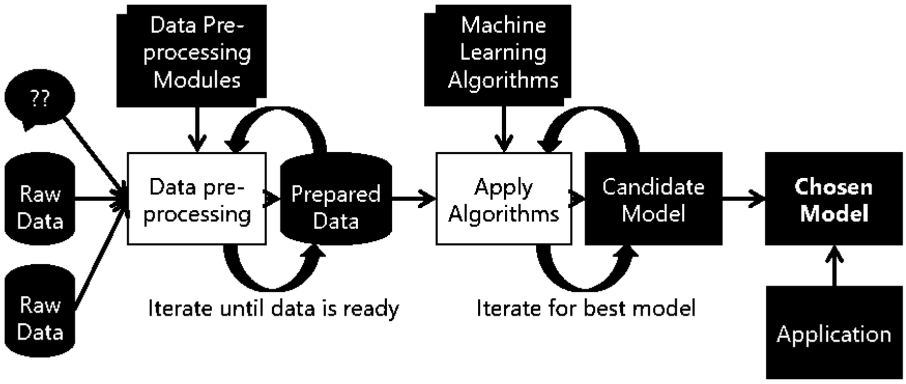

Figure 14: Machine Learning Process

Raw data flows into the system and is pre-processed in order to become prepared data. In this step, data is cleansed and malicious data are removed. This is a continuous process of new input data towards data processing towards prepared data. This prepared data is then the input for the designed algorithm. This algorithm is applied on the data in order to develop a candidate model. This model constantly improves, based on the data and the machine learning algorithm. The machine 'learns' and makes the model better. Once this model reaches its best formation, it is chosen as the final model. This model is the basis for further usage and application.

- In order to extract information from clinical or biomedical texts, a **text mining** technique is applied. Information is extracted and retrieved from text and combined to come to insights concerning certain symptoms of diseases or the effect of certain medicines that are used in combination with each other. The phenomenal growth of biomedical literature is the main reason healthcare practitioners switched to text mining analytics, as scientists need assistance in assimilating the high rate of new publications and discoveries (Zweigenbaum, Demner-Fushman, Yu, & Cohen, 2007). The purposes of this text mining approach are to recognize biological named entities, to identify relations between biomedical entities and extract valuable information from all publications and texts available. *Figure 15* shows how to go from text to insights and knowledge creation (McDonald & Kelly, 2012).

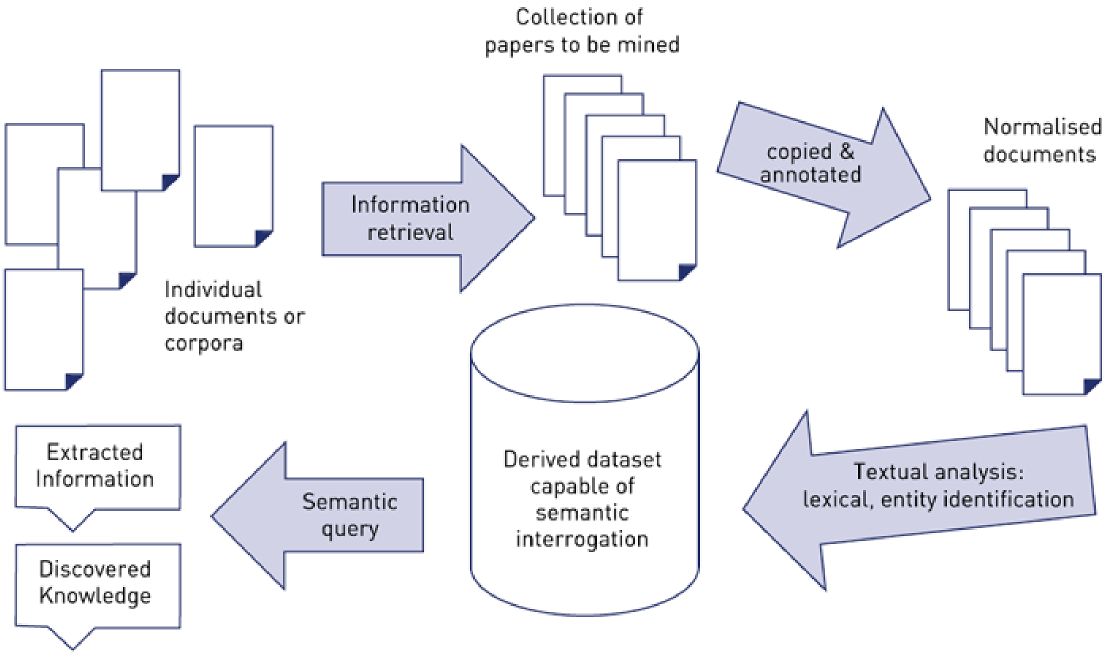

Figure 15: Text Mining Approach

This process starts with information retrieval. Individual documents of different sources are collected in order to be mined. The right information and papers are normalized via copying and annotated versions of the original document. This means they are all converted into a similar format to aid processing. Once the documents are normalized, the variety aspect of Big Data is omitted and textual analysis becomes possible. During this analysis, the entity identification recognizes the location of the names or words in the text whereas lexical identification breaks the syntax of the sentences into a series of tokens. This results in a derived dataset that is capable of semantic interrogation. Via semantic queries, information can be extracted from this database and knowledge can be discovered via certain patterns in the text that is analyzed.

5.2.2 Trends

Similar to the other discussed business domains, the healthcare sector undergoes disruptive, data-driven changes. More and more healthcare evolves towards a **new value framework**. Traditional tools do not always take complete advantage of the insight provided by Big Data Analytics, as these focus more on reducing costs rather than improving outcomes for patients. The latter should however always be given priority. More stakeholders will only benefit from Big Data Analytics insights if hospitals take a more holistic, patient-centered approach to value, one that focuses on both healthcare expenses and patient treatment outcomes equally (Datanami, 2013; Walsh, 2015). This new value framework focuses on five concepts: right living, right care, right provider, right value and right innovation (Kayyali, Knott, & Van Kuiken , 2013). Right living means encouraging patients to play an active role in their personal health by making the right choices when it comes to diet, sports, prevention and lifestyle habits. Right care means that patients should always receive the most appropriate and timely treatment available. This requires a coordinated approach with a shared information system for different caregivers working towards the same goal. This also avoids duplication of efforts and suboptimal treatment strategies. The right provider is selected based on skills and abilities and s/he must try to achieve the best outcome. Right value denotes the continuous improvement thought of professional caregivers when it comes to healthcare quality (Philipson, 2016). Right innovation refers to the individual search of stakeholders in the healthcare process for identifying new therapies and approaches towards healthcare delivery.

It is expected by the European Commission (2013) that **public expenditure on healthcare and long-term care is going to increase by 30% by 2060**. This is caused mainly by the rapidly aging population (inverted population triangle), rising prevalence of chronic diseases and costly developments in medical technology (Heinrich, et al., 2016). Thus, there is an obvious need to improve the sustainability of the current health system models and most countries are already encouraging hospitals to follow this trend. It is proven in the research of the OECD (2010) that public spending savings could approach 2% of the GDP on average by improving the productivity of the healthcare system. For Europe, this account for approximately €330 billion (Heinrich, et al., 2016). Big Data technologies are supposed to enable breakthroughs and open up new opportunities. The equal focus on quality, access and cost of healthcare is different from the traditional focus of hospitals on cost, while meeting a certain standard of caregiving. This trade-off can be optimized by using Big Data technologies. *Figure 16* gives an example of this trade-off (Heinrich, et al., 2016). This triangular trade-off is always present in healthcare, so it is extremely important for a hospital or caregiver to formulate how to address this trade-off in order to shape expectations.

Figure 16: Trade-Off Triangle Healthcare

5.2.3 Issues & Main Challenges

Based on scientific literature, we identified **four main issues/challenges** for healthcare analytics. There are of course more issues, but concentrating on these four can solve most of the (often delicate) problems. Some challenges aren't discussed here as these are a general challenge or issue for Big Data applications in an OR environment. These are discussed in *section 6.2 Future Challenges*.

- First of all, **managerial issues** are present throughout the whole process. To realize the benefits of healthcare analytics, the organization needs to establish business analytics as an organizational and cultural objective, a component of its long-term strategy (Ward, Marsolo, & Froehle, 2014). The most important concern for managers are the people that comprise the healthcare analytics process, namely patients, physicians, nurses or other medical staff – all key stakeholders in the process. Empowering these people and increasing the transparency and quality of decision making are key goals for the analytics initiative. The benefits of healthcare analytics – such as automated routine decisions, evidence-based medicine practices, real-time information – can only be realized if clinicians, supporting staff and management all understand and appreciate the importance of analytics as a tool and a fundamental process within the organization. The lack of qualified individuals and the battle to retain capable and skilled personnel that can perform these often complex analytical and data-management tasks are also two important managerial issues that need to be noted. Manyika et al. (2011) estimated in 2011 that by 2018 there could be a shortage of 140.000 people with the appropriate analytical skills in the US. The battle of healthcare organizations against other companies that benefit from Big Data is also not in favor of the healthcare sector. For the time being, it even seems that this number will be exceeded by next year.
- Another main issue that is often cited, is that of **data quality**. This problem reoccurs in many business areas, and although it is not quite pleasant that there has been some noise or mistakes or irregularities in financial data or business data, healthcare data quality is inhumanly important (Miriovsky, Shulman, & Abernethy, 2012). The need to characterize the data's fitness for use before utilizing it for any purpose in the data system is present because otherwise, there can originate a shortage of data at certain locations (because they are incorrect, outdated or not applicable) or a certain set of biases (Weiskopf & Weng, 2013). Two issues related to data quality are the price of acquiring qualitative data and the standardization of health data. Whereas the first is rather self-evident, the latter can be explained easily. When building an organizational culture that is data-driven, standardization in the processes and systems becomes the norm. In a

healthcare environment, this can lower the accuracy of electronic health data by means of misinterpretations (Ward, Marsolo, & Froehle, 2014).

- A problem of frequent occurrence is that of **data collection**. The continuous data collection process and data quality are inextricably linked, as organizations need to focus on high quality data collection (Redman, 2013). It is far better to have a smaller set of high-quality and rich data than a large set with a lot of noise and irregularities (Ward, Marsolo, & Froehle, 2014).
- A final large issue is that of **data privacy and governance**. The discussion of who is the owner of certain health information, who can use it and for which purpose is often not crystal clear. There are laws and regulations on national and European level concerning the protection of health data. American regulations are also far more different than European. Next to this, there are specific codes of conduct by occupational groups and in certain institutions that set additional requirements on how to cope with privacy-sensitive information. Several commercial parties could benefit largely from this information, so data handling needs to be conducted carefully. The different laws and regulations across national borders are also the breeding ground of many problems (Ottenheijm, 2015).
- **Other challenges** that relate to the above are heterogeneous patient sources, understanding clinical notes in the right context, the complexity of genomic data requires extra layers of technology on top of the Big Data technology and capturing the patient's behavioral data through several sensors (his/her various communications and social interactions) (Sun & Reddy, 2013).

5.3 Big Data in Operations & Supply Chain Management (O/SCM)

Over the past few decades, ICT has served as one of the most important prerequisites for successful operations and supply chain management and it is expected to do so in the future (Pahl, Voss, & Sebastian, 2017). As well the logistics as the SCM field are developing very dynamically. Think of the capacities of enterprise resource planning tools (ERP), sensor networks, social network activities, cloud applications… Big Data applications are well recognized in the field of O/SCM and offer opportunities the sector has been waiting for a long time, but they also provide challenges in handling and decision analytics. The technology is present in the industry, but many organizations still need to learn how to use it properly or how to integrate it in their existing supply chain system. Scientific algorithms and methodological developments in the field of Operations Research need to be put into running systems, as the converting of theory into practice is vital for organizations to take advantage of this research. Only then is the research considered relevant and can it have a great impact (Liu & Yi, 2017).

Until today, little attention has been given to the use of Big Data Analytics for increased information exploitation in supply chain management (Kache & Seuring, 2017). Although it is gaining increasing attention in management, empirical research on this topic is still rather scarce. The qualitative exploration is still lagging behind in literature. There is little to no consistency in defining Big Data, identifying its purpose or establishing its role in supply chain management (Richey, Morgan, Lindsey-Hall, & Adams, 2016). In their research, Richey, Morgan, Lindsey-Hall, and Adams (2016) showed that practitioners in O/SCM have distinct definitions on the concept of Big Data, depending on the context. The common factor in their findings is that almost all practitioners agree that the huge available amount of data is going to transform business technology significantly. The future perspectives however differ widely on how to handle this massive data growth and rapid technology changes.

A practical example can be found in the research paper of Matthias, Fouweather, Gregory, and Vernon (2016). A multi-channel retailer from the UK had 19 branches nationwide and consisted of three divisions: inbound, retail and re-sales. Many year's data from multiple outlets tracking all sales operations existed and the organization wanted to set up some sort of analytical platform to analyze it. A first problem that arose was the variety of the data: inconsistent, incomplete and inaccurate data were captured across the different outlets and delivery channels. Thus, the data needed to be 'cleaned' because of non-standardized data capture processes. An unexpected outcome for the company after setting up a Big Data platform and starting to analyze the data, was that certain sales patterns were discovered that nobody ever imagined, since they basically only looked at historical sales data to determine whether they were doing good business. The findings from the analysis of all company data were the breeding ground for a new framework for decision making regarding both strategic and tactical aspects of the business. The problems they faced when setting up this Big Data platform also led to the standardization of data capture across the divisions and different channels to reduce future inconsistency. The management of the company has also recognized that the platform can be integrated and used for many other operational aspects. According to them, further mining of Big Data can help them enhance its market leader position.

This example clearly demonstrates the importance of Big Data Analytics in industry. The adoption of Big Data Analytics is slowly making its way towards a massive scale-adoption. In the remainder of this chapter, we will first discuss the distinct prominent applications, after which we sketch a future perspective for O/SCM regarding Big Data Analytics. At the end, we also give an overview of frequent and important issues and challenges the industry has to cope with.

5.3.1 Applications

There are many applications of Big Data in the O/SCM domain, as this domain comprises a lot of different subdomains. We focused on the most prominent applications, and categorized them into three distinct sections. **Predictive maintenance** is one of the most important uses of Big Data in the operations world, as this combines machine data with historical data into a Big Data approach towards prevention of machine failures. The system learns about past data records by some sort of machine learning, discussed later on in this part under the techniques section. This techniques section summarizes the techniques used in the three divisions of applications in O/SCM. The second application area discussed is that of **process improvement**. The name already explains what happens here, while process data is the enabler of obtaining the best solution of the designed model. Thirdly, **risk management** is a popular application area too. Organizations try to minimize the risk during operations or try to understand it. An extensive literature study on scientific research in the O/SCM domain can be found in the Annals of Operations Research (Mishra, Gunasekaran, Papadopoulos, & Childe, 2016).

1) Predictive Maintenance

There are several techniques used and combined in maintenance decision making and optimization of multi-component systems. The common factor is the predictive nature of these techniques, making predictive Big Data Analytics particularly suitable for maintenance purposes. As outlined in the introduction of O/SCM, the complexity of industrial equipment is ever increasing. This introduces many interdependencies between the different components in the supply chain and between the machines in the different plants (Van Horenbeek & Pintelon, 2013). The predictive maintenance model should also be dynamic and real-time, because otherwise the costs of implementing and maintaining this system is greater than the benefits. The purpose of predictive maintenance models is threefold. First and foremost, failure prediction is the goal. An accurate prediction of when a machine will fail based on historical data and the current state of the machine, the operational data and performance data are combined to come up with a prevention plan. This plan ensures that certain parts of machines are replaced before the machine actually breaks down, preventing worse things to happen. Hence, maintenance and repair time is scheduled in such a way that the operational chain is interrupted as short as possible, at times it is less costly. Scheduled downtimes of the system are also way cheaper and less time consuming than unscheduled downtimes. Cost reduction is the second purpose of predictive maintenance. This is actually a logic consequence and underlying objective of the failure prediction, but machines and supply chain are approached purely from a cost perspective instead of an operational perspective. A third approach is one of improving safety. Several governmental institutions keep track of machine data and failure rates in order to continuously improve the safety of the system. *Table 17* provides a brief summary of these three purposes, along with some use cases that approach maintenance from a different perspective.

Goal	**Failure Prediction**	**Cost Reduction**	**Improve Safety**
Use Cases	Railways, MTConnect, Augury	Arimo, PTC Case Study	Siemens Tracksure
Data	Failure data, machine data, maintenance history data, sensor data	Machine data, condition data, sensor data, failure data	Sensor data, machine data, failure data
Papers	(Adomavicius, 2014), (Li, et al., 2014), (Nelson P. , 2015)	(PTC, 2015), (Tran & Pham, 2016)	(Tyler, 2016)

Table 17: Predictive Maintenance Summary

2) Process Improvement

Process control (and thus improvement as well) has been around in the industry for over twenty years, but recently it has taken a new direction. The rapid evolution of Big Data solutions facilitated significant advancements in process control by providing step change improvements in capabilities. Real time analysis and process control/monitoring is becoming the standard throughout the industry. In the PhD research study of Liedtke (2016), it has been argued that a company can benefit from integrating Analytics and Big Data with quality. The application of quality principles can potentially improve Big Data and Analytics initiatives and this can result in new sources of customer values or even a new source of competitive advantage for companies. The continuous drive from organizations towards process innovation calls for new technologies, offered by Big Data. There is however one very important remark on this notice: operations guys or data scientists analyzing the process data need to know and understand the process thoroughly and truly understand the flow of operations. Only then it becomes possible to build and adjust the models towards process improvement (Dutcher, 2014). The purpose of using Big Data techniques in a process environment is to continuously innovate and to control the quality of the processes (in line with the total quality management philosophy). The following five key impacts of Big Data and its applications on quality management are opportunist for O/SCM (Roberts, 2014):

1) Correlating performance metrics across multiple plants
2) Perform predictive modeling of manufacturing data (cfr. the previous paragraph)
3) Better understanding of the supplier network performance
4) Faster customer service and support
5) Real-time alerts based on manufacturing data (leans towards the fault detection purpose)

Another purpose that serves process improvement is fault detection. This mechanism aims to identify defective states and conditions within industrial systems, subsystems and components. Since these systems tend to be high-dimensional in measurement nowadays, the traditional approach towards fault detection is no longer self-evident (Zhang L. , 2016). The system should support the capture and real-time analysis of hierarchical process data, allowing the analysis to take place at multiple levels (both organizational and process levels). A final purpose of process improvement is to eliminate waste from the entire process and by extension the entire supply chain. By crunching huge numbers, Big Data Analytics enables companies to have an optimal stock policy and an efficient, leaner supply chain. This eliminates waste in all components. *Table 18* provides a concise summary of process improvements, along with some use cases.

Goal	**Process Innovation**	**Quality Assurance**	**Fault Detection**	**Eliminate Waste**
Use Cases	Call Center Case Study, Biopharmaceutical Case Study, Beetrack	Intel	IBM, Fuji Electric[27]	IBM, Datameer[28]
Data	Process data, component data, historical data	Component data, process data	Plant state data, log data, process data	Plant state data, log data, process data, SC components data
Papers	(Auschitzky, Hammer, & Rajagopaul, 2014), (Honorato, 2016), (Vera-	(Bertolucci, 2013)	(Ittmann, 2015)	(Dhawan, Singh, & Tuteja, 2014), (Ittmann, 2015)

[27] https://www.fujielectric.com/company/research_development/theme/bigdata_plant.html
[28] https://www.slideshare.net/Datameer/lean-production-meets-big-data-a-next-generation-use-case

Baquero, Colomo-Palacios, Molloy, & Elbattah, 2015)			

Table 18: Process Improvement Summary

3) <u>Risk Management</u>

No matter what happens in the supply chain, upstream or downstream, every manufacturer or component of the supply chain will try to minimize its own risk (Chopra & Sodhi, 2014). This often contravenes with the optimal global solution of the supply chain, but it is a natural behavior that a manufacturer tries to achieve its local optimum. Risk can take many forms and thus applications are very divergent as well. Several risk models have originated ranging from predicting probabilities of delay to lower customer demand by incorporating several market risk factors. All of them make use of Big Data to better predict future outcomes and also prescribe what actions to take, taking into account this risk. The rapid analysis of different information can assist organizations to forecast events that influence the supply chain and thus enables them to take measures for minimizing the associated potential risk. This includes both risk evaluation as resilience planning. The continuous monitoring of the entire supply chain facilitates the detection of performance deviations. The risk analysis process makes use of Big Data Analytics in all of its steps: organizations first identify the risk, then asses it after which they develop the right responses to the risk. The latter is closely related to the decision-making process in other domains. The final step is to develop a contingency plan or to take preventative measures for the risk in order to minimize the total risk. *Table 19* provides a succinct summary of the applications of Big Data in risk management. Use cases are provided for the reader who wishes to further exploit this topic.

Goal	**Identify risk**	**Reduce/allocate risk**
Use Cases	Citigroup, DHL, Oracle, SAP	Deloitte, Capers Jones RiskMaster
Data	Social media data, supply chain data, supplier data, customer data, performance data	Projects data, historical data, process data
Papers	(Ittmann, 2015), (Nash, 2012), (Olson, 2015), (Stackowiak, Mantha, & Licht, 2016)	(Davis, Nayeri, Suryanarayan, & Wilson, 2015), (Nash, 2012), (Reitano, 2011)

Table 19: Risk Management Summary

4) Techniques

There are several techniques used in Big Data applications in the O/SCM field. Some of them are already discussed in previous sections, so we focus here on techniques that aren't delineated yet and that are very often used in the field.

- To find the root cause(s) of a problem that occurs when preventive maintenance wasn't satisfactory, a **causal analysis** (or root cause analysis) is designated. In a causal analysis, the independent variables are considered as causes of the dependent variable. The aim of the analysis is to determine whether a certain independent variable really affects the dependent variable and what the magnitude of that effect is, if any. A root cause is the basic reason why something happens and this needs to be taken care of, not one of the symptoms it causes. The root cause analysis tree diagram (*figure 17*) indicates the breakdown from problem to corrective action. Big Data and Analytics are used in order to find the root cause. Starting from the problem, the causes are visually mapped. The main reason of these causes (= root causes) can be found with evidence-based techniques or simply by interviewing the personnel. The latter will then reveal their insights on the problem. Once these root causes are identified, corrective actions can be taken. It is important to adjust this action to the root cause and not simply to a cause of the problem. Otherwise, latent problems will remain present in the organization.

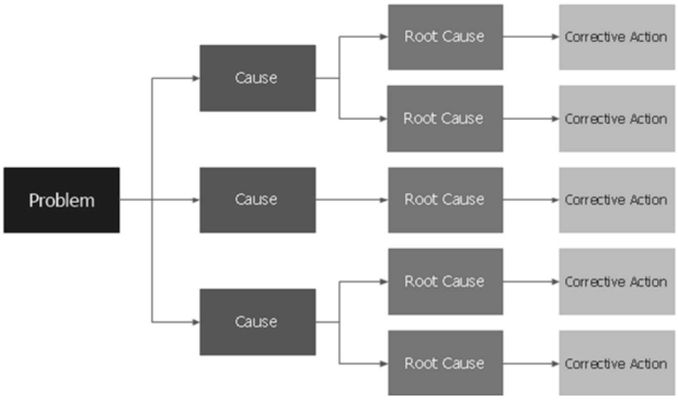

Figure 17: Root Cause Analysis Tree Diagram

- **Data mining** follows more or less the same approach as text mining (*see section 5.2.1 Applications of Healthcare*), but differs from it because it is even more comprehensive. It is the computational process of discovering several patterns in large datasets involving different methods such as artificial intelligence, machine learning, statistics… The purpose of data mining is to bring out knowledge and significant information that is hidden in the combination of available data (Toma, 2010). Trends in the data can be identified as well. The scope of data mining in O/SCM reaches from anticipating the future demand (forecasting), to the optimal supplier base (procurement), to the optimization of transportation moves, to monitoring the state of production and the supply chain ecosystem (risk management), to logistics optimizations and to the analysis of overall patterns influencing the planning and managing of the daily operations and the supply chain in general (Lehmacher, 2016).

- The field of study interested in the development of computer algorithms for transforming data into intelligent actions, is called **machine learning**. Massive data growth necessitated additional computing power, which in turn inspired the development of statistical methods for analyzing these larger datasets. This triangular relationship, depicted in figure 18, is the foundation of machine learning (Lantz, 2013a). The more available data, the more advanced the statistical methods and the more computing power. These three are very strongly interrelated and this continuous cycle of advancement allows even larger and more interesting data analyses. The essence of this technique is to make sense of complex data, in a desired time span. Machines are said to 'learn' if they are able to take their previous experience and utilize it in such a way that it improves their future performance on similar experiences. The five steps of the machine learning process are collecting data, exploring and preparing the data, training a model on the data, evaluating model performance and improving the model performance (Lantz, 2013b).

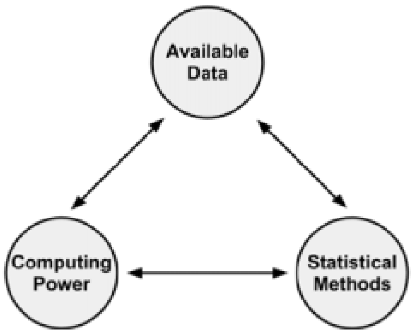

Figure 18: Triangular Relationship Machine Learning

5.3.2 Trends

There are many trends noticeable in the O/SCM field, as this is a continuously improving field. Many firms and organizations exhibit the same changing characteristics towards a better outcome, resulting in many widely-adopted trends. Organizations look at the industry and try to imitate competitors or large firms that perform well, but all firms try to be unique in a way of conducting this change or in riding these waves/trends.

The trend towards **incrementalism** in the O/SCM profession has led to the challenge of practitioners to demonstrate the worth of O/SCM to business and society. However, to get published in the better journals, researchers may not be too innovative in their thoughts or even with their methodology

(Simpson, et al., 2015). This causes research in the O/SCM field to be rather incremental than disruptive and of low relevance. There is a great gap between industry/practice and theory and to be directly useful, research needs to be not too innovatory because otherwise, the industry can't keep up and the research paper gets rejected or given less attention. This is a notable trend, due to the inertia of many big companies.

Firms mainly focus to become an expert in a particular field and specialize in one particular process, product or component of the supply chain. This of course creates a much more complex entire supply chain system, requiring **more flexibility** from the firms involved. Companies try to be more flexible and more apt to change by using Big Data Analytics in their Operations Research approach. The ability to adapt rapidly to changes in the product mix demands, in volumes required by a successor in the supply chain or in delivery schedules has become a major competitive strategy. It has also become a major competitive advantage for firms that showcase the highest extent of flexibility. Some authors refer to this as agile manufacturing (Dubey & Gunasekaran, 2014). More flexibility comes at a certain cost, so Big Data Analytics are applied in order to cut costs and find appealing insights at a high pace. Because speed and insights can lead to more flexibility.

A third noticeable trend is the **rise of data-driven business process re-engineering** (BPR). BPR involves drastic measures or break-through improvements to increase the performance of a firm. Operations managers must look at the existing processes and how they are interrelated. Then they have to analyze this as-is model and try to figure out whether and where there is a bottleneck or where operational improvements are possible. Big Data Analytics can help them to come to these insights. The principles of methodologies like Six Sigma, Kaizen and the Theory of Constraints are considered when redesigning the business process map. According to a Deloitte study (2013), the integration of an organization's three key dimensions is critical during process re-design. People, processes and technology must go from intersected to integrated, as *figure 19* shows (Deloitte, 2013). A clear governance structure as support is critical in sustaining a data-driven and fact-based organizational culture and the effective capture, analysis and movement of data.

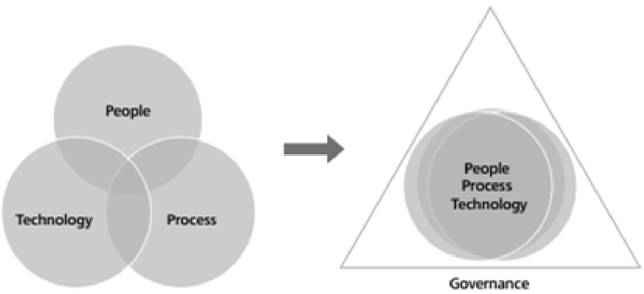

Figure 19: From Intersection to Integration of Key Dimensions

5.3.3 Issues & Main Challenges

As often with new, disruptive applications in business domains, the benefits come at a certain cost and with certain challenges. For O/SCM, this is not different. Most difficulties are already present in traditional approaches, so these are not inextricably linked to the Big Data approach, but more to the nature of O/SCM. The recent overload of data and the opportunity to gather and monitor more data at the same time, do make sure that the capacity of traditional systems need to be revised and that these new systems need to cope with the following challenges.

- Management decisions that are based on the insights given by Big Data Analytics are only as good as the data on which they are based (Hazen, Boone, Ezell, & Jones-Farmer, 2014). The **data quality problem** is tragically strong present in the O/SCM environment. Many methods are thought out for monitoring and controlling data quality, resulting in a total quality management approach inside O/SCM. In the past, many authors have shown the impact of poor data quality on business decisions (Warth, Kaiser, & Kügler, 2011) and even estimated the losses to exceed billions of dollars per year (Dey & Kumar, 2010). Next to this, there are intangible losses connected to the poor data quality as well such as customer dissatisfaction, suboptimal decisions, job dissatisfaction and propagation of mistrust between and within organizations.
- **Complexity** derived from the physical supply chain network, the product and service mix and the systems that manage the supply chain propagate in the complexity of coming to Big Data Analytics insights. Big Data Analytics however are the solution to the exponentially growing complexity in O/SCM, as internationalization, increased collaboration and growing volumes of different products and volumes are key characteristics of the second decade of the 21st century (Leveling, Edelbrock, & Otto, 2014). To increase supply chain and operational visibility to handle complexity and to support decision making, Big Data technologies will need to keep evolving in order to cope with this increasing complexity. Model management is also an issue related to the complexity of today's supply chains. Severely time-compressed decision loops need to run thousands of models in just a few seconds or minutes, which implies that models need to be very up-to-date and disruptive to handle this complexity.
- At the moment, the **relevancy of research** in the O/SCM domain is quite lagging the disruptive technology developments. The link with managerial practice is often disregarded. It is rather a tactical domain, whereas business school focus more on strategic issues, leaving operations behind. There clearly is a growing research gap in this domain, but *section 6.1 Research Gaps* is devoted to this topic.

5.4 Big Data in Public Services

Since the financial crisis struck many organizations, banks and governments in 2007-2008, a lot has changed in the pursuit of doing business and the development by public services and the public sector in general. The main impact of the crisis was that cost improvements and high efficiency became the leitmotif in shaping public services and in innovative public projects (Curry, Blijleven, & Van de Walle, 2014). This may of course not affect the quality of the services, so new technologies and approaches needed to be considered. The financial crisis became an incentive for governments and public authorities worldwide to innovate towards cost-efficient, high-efficiency and qualitative public services for its citizens. Some countries and authorities however experienced this reformation as rather negative, with an adverse effect on services (Claessens & Kodres, 2014). Worldwide, increased attention was given to performance-related reformations, along with media attention and increased citizen expectations. This enhanced the transparency and openness of the public sector processes and gained more trust and comfort by citizens who make use of the services. On the other side, the financial crisis ensured a rupture in citizen trust in government, which is now still being restored. Many authorities and organizations became privatized as well (Curry, Blijleven, & Van de Walle, 2014).

With the raise of Big Data and its tendency towards a closer relationship with the Operations Research field, discovering new ways for analyzing and managing Big Data to create value would also increase the accuracy of predictions in predictive models. This improves the management and security of transportation infrastructure and enables more informed decision-making in the entire transportation industry (Vlahogianni, Park, & van Lint, 2015). These accompanied challenges may drive new opportunities or insights and even transform the way transportation and traffic engineering phenomena are perceived nowadays. Public authorities often dispose of exuberant amounts of data, but most of them are never used[29]. The adoption of Big Data Analytics in public services can help to crunch these often unused data to help shape the urban landscape and find ways to cope with problems related to public transport and traffic.

Another challenge public services and governments face, is the increasing population living in the city. This puts pressure on environmental issues, livability, housing, transport… Cities need to become 'smarter' in order to keep carrying this increasing population. A great example can be found in the city of London (Weinstein, 2015). London is the second largest city in Europe and one of the largest cities in the world. Therefore, a large and comprehensive public infrastructure is required to make this city livable. A reasonable amount of the governmental budget goes to public transport, as this is considered by the

[29] https://www.accenture.com/us-en/insight-highlights-public-service-big-data-big-gains

English authorities to be the main differentiator for a city to make it livable. To give an example of the magnitude: in 2014, nearly 2.5 billion journeys were made by bus and 1.26 billon by the London Underground. With a growing number of citizens expected, the city is faced with these public service challenges and they are urging. The city of London makes use of Big Data and Analytics for numerous things in transport: it stores the Oyster and contactless public transport cards, it analyzes bus location data, traffic information is shared real-time, social media are used to communicate with travelers, traffic capacity planning, automatic fare refunds, influencing travel… (Weinstein, 2015). Big Data Analytics is seen as the answer to many questions and challenges the city faces. This brings us to the next section of Big Data Analytics in public services, namely its applications.

5.4.1 Applications

The above introduction of Big Data and Analytics in public services already suggests there are many applications of Big Data Analytics in the public sector and where governments can make use of. All serve a particular purpose or are used to cope with a certain challenge. Many challenges are addressed simultaneously as well, since they are often interrelated. First of all, **smart cities** are discussed. Smart cities serve different purposes and is a first category for Big Data applications. Challenges such as transportation, housing and waste management are considered in smart cities. Secondly, **public data hubs** are laid out. More and more, enormous data amounts are gathered in one place, accessible by governments and public services. Often times, the government opens (parts of) these hubs to private companies, in order to stimulate the economy. The purposes of these data hubs are thus to improve transparency and to personalize citizen experiences, but is also used on a large scale to reduce tax and social security fraud. Thirdly, the application area of **security intelligence** is touched. In uncertain times, Big Data Analytics can be used to counteract terrorism and to employ efficient law enforcement, but also to raise regulatory compliance. Subsequently, the techniques used in these applications are discussed. This gives the reader a clear insight in the process these applications go through.

1) Smart Cities

As the concept of a ‘smart city’ gains wider currency, there is still confusion about what it really is. The importance of smart cities is discussed in scientific literature (Mori & Christodoulou, 2012), as cities play a prime role in social and economic aspects worldwide and they have a huge impact on the environment. As an illustration: 75% of the population in Europe lives in urban areas and the number is expected to increase to 80% in 2020 (Albino, Berardi, & Dangelico, 2015). Urban sustainability in all its aspects is on top of the agenda for almost all countries in the world. The promotion of an anthropocentric approach, according to which cities should respond to people’s need through sustainable solutions for social and economic aspects, is obliged according to many authors (Berardi, 2013a, 2013b; Turcu, 2013). What a smart city looks like is summarized in the following table (based on Lombardi, Giordano, Farouh, & Yousef, 2012):

Components	Related aspect of urban life
Smart economy	Industry
Smart people	Education
Smart governance	e-democracy
Smart mobility	Logistics & infrastructure
Smart environment	Efficiency & sustainability
Smart living	Security & quality

Table 20: Smart City Components

The six components in *table 20* are more or less the same as the purposes of smart cities, namely improving these components as regards costs, efficiency and sustainability. Smart economy relates to the industry and the way firms do business in a smart city. Cities and governments work more closely with companies in the city in order to create synergies, based on shared knowledge and data. Smart people relates to the education system. Education is based more on technology and the internet, as young people are the future workforce of this digitalized world. More and more integration of internet and software is included in the smart education system. Smart governance relates to how governments and public services govern their services. Smart mobility on its turn relates to public transport, traffic engineering, smarter cars etc. A smart environment is related to waste reduction in cities as regards energy consumption, garbage collection and public parks/domains. Smart living is the final component and relates to the security of citizens in smart cities and their quality of living. The following tables (*table 21 & 22*) summarize the applications of Big Data (Analytics) that help cities to become smarter. Use cases are provided for the interested readers who want to know more about certain aspects of smart cities. This topic is very extensive and is a very hot topic in scientific literature, so this summary is rather to give a clear and concise idea than it is a complete collection of use cases and papers with extra information.

Goal	Smarter Economy	Smarter People	Smarter Governance
Use Cases	Power of Three (EY), APIs	Smart Schools, Sandy High School[30], Grass Valley Elementary School[31]	SCiGov[32], Neighborland[33]
Data	Population data, sensor data, open databases	Population data, student data, school data	Population data, administrative data, infrastructure data, social media
Papers	(Atalla, Banks, Littlejohn, & Hiscock-Croft, 2016), (Pettey, 2016)	(Williamson, 2015)	(Kumar, Singh, & Gupta, 2016)

Table 21: Smart Cities Summary (a)

Goal	Smarter Mobility	Smarter Environment	Smarter Living
Use Cases	Transport for London, Uber, Lyft, Zipcar, Car2go, Carma, Zimride	Forest fire detection, earthquake early detection, monitoring air/water/soil quality, Padova Smart City	Ambient Assisted Living (AAL), Keep In Touch (KIT)
Data	Population data, traffic data, GPS/location data, smart camera data, sensor data	Population data, weather data, GPS/location data, sensor data	Population data, health data, sensor data
Papers	(Viechnicki, Khuperkar, Fishman, & Eggers, 2015), (Vlahogianni, Park, & van Lint, 2015), (Weinstein, 2015)	(Rashidi, Cook, Holder, & Schmitter-Edgecombe, 2011), (Zanella, Bui, Castellani, Vangelista, & Zorzi, 2014)	(Dohr, Modre-Osprian, Drobics, Hayn, & Schreier, 2010), (Hmida & Braun, 2016)

Table 22: Smart Cities Summary (b)

[30] http://www.ibigroup.com/projects/sandy-high-school
[31] http://www.ibigroup.com/projects/grass-valley-elementary-school
[32] http://smartcitygovernance.eu/
[33] https://neighborland.com/

2) Public Data Hubs

Governments around the world see Big Data Analytics as an opportunity to help them meet their goals in different directions. The free and continuously enlarging flow of information from public organizations/governments to citizens promotes greater trust between these citizens and their government. This improved transparency creates a mutual trust between citizens and governments and results in richer data hubs. It can also be used to collect, organize and analyze large amounts of data from government computer networks to give cyber defenders greater ability to detect and counter malicious attacks. Governments also use Big Data Analytics to better understand national or regional sentiment or to learn what citizens need. Public services and the addition of these are adapted to those needs. If there is some agitation to a certain extent in a city or district, adjusted police forces can be put to work. Tax and social security fraud can be more easily detected as well in large amounts of data. Algorithms are used to detect patterns and thus find suspicious transactions occurring in real-time. Tax paying behavior is also mapped with the help of different data, at local and national level. *Table 23* gives a small summary of public data hubs and its goals, accompanied with some use cases.

Goal	**Improved Transparency**	**Personalized Citizen Experiences**	**Reduce Tax & Social Security Fraud**
Use Cases	Data.gov[34], Open Government Data[35], Open Data Portaal[36], Chicago Open Data, DataMarket[37]	Smart Tourism Destinations, hyper-personalization	Internal Revenue Service
Data	Population data, public available data	Population data, cookies data, public data, social media data	Population data, social data, local and national data sets
Papers	(Kassen, 2013)	(Celdrán Bernabeu, Mazón López, Giner Sánchez, & Ivars Baidal, 2016), (de Roys, et al., 2017), (Lohrmann, 2014)	(Kim, Trimi, & Chung, 2014)

Table 23: Public Data Hubs Summary

[34] https://www.data.gov/
[35] https://opengovernmentdata.org/
[36] https://data.stad.gent/
[37] https://datamarket.com/

3) Security Intelligence

At times of great uncertainty, security of citizens is of the highest importance in many countries. Big Data Analytics can be used to enhance this security. First of all, federal governments use it to combat terrorism. Data lakes with both classified and unclassified information are explored in order to identify threat patterns and predict potential sources of domestic terrorism. To control access to these lakes and to protect personal privacy, coded data tags are connected to the data logs such that this information does not fall into the wrong hands and privacy is always guaranteed. A second purpose of security intelligence is to efficiently deploy police/law enforcement. By analyzing crime trends and historical data, Big Data Analytics can have a significant impact on law enforcement. Relationships are assessed between people by analyzing places and factors such as age categories, social standards… By crunching these numbers, crimes are predicted in neighborhoods and what the chances are something is going to happen. More officers will patrol in these neighborhoods when a crime is likely to happen. This can either prevent crimes from happening or prevent further escalation by the proximity of law enforcements. Security intelligence is also ideal to improve safety in general. Online data and cookies are used in order to find patterns of fraudulent practices or security breaches. Regulatory compliance is the main purpose here. *Table 24* provides a small summary on the subject of Security Intelligence.

Goal	**Counteract Terrorism**	**Efficient Law Enforcement**	**Regulatory Compliance**
Use Cases	GIS Federal[38], Neptune & Cerberus (Homeland Security)	Los Angeles Police, Durham Police Department North Carolina	ECHO[39], PRISM, NSA Google Cookies
Data	Surveillance data, sensor data, GPS/location data, social media data	Crime history data, social media data, census data	Population data, social media data, cookies
Papers	(Shi, 2014), (Strohm, 2016)	(Berg, 2014), (IBM, 2015)	(Greenwald & MacAskill, 2013), (Feloni, 2013)

Table 24: Security Intelligence Summary

4) Techniques

Of the techniques that aren't described yet, the following are used on a regular basis in security-related intelligence applications:

[38] http://www.gisfederal.com/
[39] https://echo.epa.gov/

- **Natural Language Processing (NLP)** is a subfield of Artificial Intelligence (AI), concerned with making computers understand the sayings or words written in human languages. It is a way for computers to analyze, understand and give meaning to human language in a smart and useful way[40]. Several tasks are possible such as translation, automatic summarization, named entity recognition, sentiment analysis and speech recognition. The latter two are particularly useful in a security intelligence setting. Computers can also make sense out of social media posts, blogs and obscure forums in order to prevent crime from happening or to interact when needed (Kiser, 2016). Suspicious users are further investigated and the algorithm links their social media accounts, blog profiles and forum profiles. In general, NLP follows five subsequent steps, as *figure 20* shows. First of all, a lexical analysis takes place. This involves identifying and analyzing the structure of the words. This step divides the whole text/recording into paragraphs, sentences and words. Secondly, these paragraphs are parsed in a syntactic analysis. In this step, the words in the sentence are analyzed for grammar purposes and to determine the relationship among the words. The third step is a semantic analysis. This means that the text is checked for its meaningfulness. Syntactic structures and objects are mapped in the task domain. After this, a disclosure integration is obtained by checking the sentences that proceed and succeed one another. The meaning of a sentence depends upon the meaning of the previous sentence and vice versa. A final step consists of a pragmatic analysis. Everything what was said is re-interpreted on what it actually means and whether this makes sense. Real-world knowledge is required in this final step and it serves as some sort of feedback mechanism.

Figure 20: Subsequent Steps Natural Language Processing (NLP)

- **Cognitive computing** is a little more comprehensive than NLP. It describes technology platforms that combine reasoning with machine learning, NLP, speech, vision and human interaction. The purpose of cognitive computing is to improve human decision making. The result is a computer that can mimic the way the human brain works. This is a very advanced and complex technique that is up until now not used on a large scale. Rather, there are some smaller experiments and research towards the subject. As computers will become more able to think like humans in the

[40] https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_natural_language_processing.htm

future, this will expand our capabilities and knowledge. Governments can benefit widely from this knowledge and share it with their people (Marr, 2016d).

5.4.2 Trends

Reviewing Big Data projects and initiatives in the public sector of several countries worldwide, it became clear that there are three notable trends: smart cities are slowly becoming a reality in many countries, governmental applications are rather marginally classified as 'Big Data' but this is improving and there is a rise in smart governments, called e-governments.

- Smart cities are already discussed in the applications of this sector, but over the next decade it is likely that the **real-time smart city will become a reality** in many cities. Cities are considered 'engines of economic growth' as they attract people, activities, investments, organizations and this contributes to the national economic output. It is estimated that this results in further urbanization until 2050 in almost all major regions worldwide (United Nations, 2014). From a sustainability and ecological viewpoint, this is practically infeasible. To reach the ideal state of a smart city, a closer collaboration between governments, the private sector, civil society organizations and citizens is required. The transition of a conventional urban economy to a smart economy is already gradually progressing in many cities, but this goes rather slow (Vinod Kumar & Dahiya, 2016). It is expected that this will accelerate in the near future, as technology adoption and habituation accelerates as well.
- Most projects that are operational or being implemented today can only **marginally be classified as a 'Big Data' application**. The majority of 'Big Data'-projects make use of structured databases and not of real-time and unstructured data. It is called a Big Data-project because of the high volume of the data, but this only accounts for one of the three V's of Big Data (see *section 2.1.1 Big Data*). By definition, they are Big Data projects, but they do not take on the challenge of variety and velocity. It is expected that this will increase in the future, as public services will become more familiar with Big Data technologies (Manzoni, 2017). But at this point in time, we can say that Big Data applications are still at an early stage of development in the public sector of most countries, with only a handful of projects in operation (Kim, Trimi, & Chung, 2014).
- In the next decade, experts foresee a **rise in e-governments**. Governments that are experimenting with several Big Data techniques improve governmental performances with Analytics and bring data-driven decision-making to frontline employees. Around the globe, more and more governments will start to develop such a smart government, classified as an e-government (van Rijmenam, 2015). A consequence that is closely related to this trend is the tendency of

governments to open up their data sets and work with open APIs to enable the economy (startups, enterprises, public organizations…) to easily connect with their government. This creates a synergy, as more data becomes available for the government, which benefits future public services, and the economy of the country benefits from the exposure of data.

5.4.3 Issues & Main Challenges

Governments often dispose of even more data than private companies, but this brings along more complex issues and challenges. Who bears the responsibility of all those data? How can a government make sure that there are no security breaches and that privacy is guaranteed? Is all data available useful? Those are but a few questions that are stringent for governments that want to become more digitalized, evolving towards an e-government. Specific challenges related to this are the following.

- With **data ownership** comes great responsibility for its management, storage, use and misuse (Morabito, 2015). But who really bears the responsibility in the age of open and public data? In theory, public organizations and governments are the custodians of our data, but they can use it like they want in exchange for providing citizens with public services and promoting public goods (Gudipati, Rao, Mohan, & Gajja, 2013).
- **Privacy** will be largely discussed in *section 6.2 Future Challenges*, but deserves to be mentioned with the challenges in public services as well. Citizens are keen on their privacy and would like to have something to say in what happens with their data. The point where that is possible is unfortunately long gone, as everything and everyone creates data nowadays and most often this is tracked as well. This however does not mean that governments can do what they want with these data. In order to maintain citizens' privacy, data encrypting and anonymization of the used data is of high importance. Governments will have to prove what value Big Data can create and why they have to collect data from citizens, for what purpose. If this is not clear, people will feel being spied upon, as if a big brother is watching them. This may of course not happen, because then data collection will become far more difficult and more issues will rise.
- Related to the previous issue, is the one of **data security**. This is also a challenge that is applicable in almost all domains, but is even more applicable for public services and governments. A security breach in of the secure data lakes from the government can have disastrous consequences as this information is highly confidential. In the wrong hands, this can be very dangerous.

The above described issues are specific to the use of Big Data and analytics in the public sector, but these are also common issues in Big Data Analytics in general. In the following chapter, a large section is devoted to the future challenges of Big Data (Analytics) in general and the incorporation in OR.

6. Future Perspective

To fully capture the whole story of Big Data and Analytics in an OR environment, it is important to not only look at how things are at the moment and which applications are already present. A glance at the future can reveal many insights on top of the established knowledge and technologies. To many people, the use of Big Data and Big Data Analytics may already seem a little futuristic, but it is not only a here and now phenomenon. You can say it is more of a journey than it is a destination. What is now possible and accessible for some organizations, will become self-evident in a few years and will be adopted on a large scale in many companies. But these applications are often only just scratching the surface of what's possible (Teradata, 2017). If we should compare the road of Big Data Analytics with a marathon, it is still in the warm-up stage, getting ready for what's further down the road.

In *section 6.1*, scientific literature is assessed. More specifically, we look at what has not been written yet or where there exist research gaps. A lot has been written on Big Data and Analytics but most papers are only marginally contributing to the general knowledge. Many calls have been made by renowned literature providers and publishers on some very specific topics related to Big Data and Analytics. This also shows the importance and urgency on further research or deeper exploitation on certain topics that possibly have high value for several business domains. The distinction is made between research gaps of Big Data and Analytics on the one hand and the combination with OR/MS on the other.

After that section, we dive into the challenges Big Data and Analytics will face in the future in *section 6.2.* Many of the challenges and issues are already discussed in *Chapter 5*, but this section will rather delineate the universal challenges and issues. Privacy, security and ethics are very interrelated and will probably be an all-time problem, whereas technology and data characteristics are inherent to the novelty and fast evolution of Big Data Analytics. Challenges inherent to the nature of OR close this section.

Section 6.3 discusses the future opportunities for Big Data Analytics and which direction it could take to maximize economic value. There are three major fields related to Big Data Analytics that could open many opportunities in the future. The first is artificial intelligence (AI). In combination with Big Data, AI creates smart systems that will think for us and that take on the challenges related to Big Data. The second one is the combination of Big Data Analytics with virtual and augmented reality. This should enable managers and policy makers to fully grasp the problem analysis by visualizing the data such that it fits in the human perception and capabilities. A third major opportunity lies in the combination of Big Data Analytics with Internet of Things. Commonplace objects tend to become smarter in the future, as they are more and more connected to the internet via sensors, buttons, cameras… This creates huge opportunities in daily management, massive-scale analyses and autonomous decision-making objects.

6.1 Research Gaps

In the following section, the research gaps of scientific literature on Big Data and Analytics are revealed. After that, the same is done for OR/MS with regards to Big Data and Analytics. The main objective in research in the near future should seek to both incorporate the unique aspects of the OR/MS discipline, as well as the innovations, concerns and unique characteristics of the Analytics age. The combination of the two disciplines has been discussed widely, but more in an analytical viewpoint than an operational.

6.1.1 Gaps in Big Data (Analytics)

Saying that a lot has been written on topics related to Big Data (Analytics) is really an understatement. Scientific literature is overwhelmed by the presence of published articles, millions of new blogposts come out every day, entire books are published on a daily basis… This is closely related to the 'hotness' of Big Data, but also reveals that there is a lot to investigate related to this topic. It also means that not everything has been written yet. During the writing of this dissertation, we experienced first-hands how difficult it was to find appropriate sources or research on certain topics.

The main reason for this dissertation was to give an overview of applications in certain business domains, which are similar to and can be extended to almost all business domains. Clearly, up until today there is no literature overview of applications of Big Data Analytics in OR. On the website of Elsevier, one of the world's major providers of scientific, technical and medical information, a special issue has been calling for the submission of papers[41]. The aim of this special issue is to collect state-of-the-art research findings on the **latest development, up-to-date issues and challenges** in the field of Big Data Analytics for business intelligence (Liang, Shen, & Guo, 2016).

Next to the call for papers of Elsevier, there has been a call for papers for the 6th IEEE International Congress on Big Data[42]. Their call especially reaches out to submissions that describe **techniques and tools for handling big variety and high veracity** as well as **Big Data applications** for enterprise, government and society. Other topics they would like to acquire are value improvement by Big Data, Big Data models and algorithms, Big Data architectures and management. Security and privacy are also very high on the wish-list. Other recent call-ups in the same range have been made by the MDPI[43], the Multi-Conference on Computer Science and Information Systems[44] and by IEEE for the International Conference on Big Data[45].

---

[41] https://www.journals.elsevier.com/expert-systems-with-applications/call-for-papers/special-issue-big-data-analytics-for-business-intelligence
[42] http://www.ieeebigdata.org/2017/cfp.html
[43] http://www.mdpi.com/journal/information/special_issues/data_driven_science
[44] http://bigdaci.org/call-for-papers/
[45] http://cci.drexel.edu/bigdata/bigdata2017/CallPapers.html

The common factor of these calls is to gather relevant sources and to give a clear overview of Big Data applications and issues related to the use of these applications. Once this is mapped out to a certain extent, best practices of these applications and techniques can be combined to form newer and even better applications. Insights can be gained from this extensive overview.

Another prevailing question in the field is **how the risk of sharing datasets should be accounted** when it is unknown what auxiliary datasets they will be combined with in the future. This is chiefly relevant for public datasets and shared datasets between the government and private companies. It is the strength and peril of large-scale data analytics that it has the capacity to combine datasets relatively easily from highly different contexts. This enables the data to become very flexible and perpetually available for repurposing, but this also brings along a major risk. These unknown and unpredictable auxiliary datasets can be used to re-identify personal data in a research dataset, which means that the risks faced by subjects of data research are not limited to the context and lifespan of a sole project. Historically, researchers have not been held accountable for such far-reaching consequences (Metcalf, Keller, & Boyd, 2016). Future research focused on empirical measurement and consistent accounting of such risks is strongly needed into scientific literature.

An unaddressed, yet important issue is that of the **ecological and environmental impact of the rise in Big Data research and Big Data industry**. The propagation of data intensive internet applications strongly drives energy usage outside of the control of industry leaders (Metcalf, Keller, & Boyd, 2016). Research towards the sustainability of Big Data in the industry is lacking in scientific literature. The trade-off between performance and energy-efficiency has not been exploited yet, as the focus now is merely on performance. The reason behind this is the focus on real-time performance and high pay-off, because the value of Big Data Analytics still needs to be proven in many organizations. Therefore, the ecological and environmental consequences are often neglected in this merging phase of Big Data.

6.1.2 Gaps in OR/MS with regards to Big Data Analytics

Mortenson, Doherty, & Robinson (2015) delineate **six future research agendas for OR/MS** in the Analytics age: leveraging Big Data volumes, utilizing new data architectures, incorporating unstructured data, streaming data and real-time analysis, visualizing data and OR/MS and the wider ecosystem. These six agendas perfectly summarize the research gaps that exist for smoother incorporation of OR/MS in Big Data Analytics or the other way around.

The first topic on the future research agenda is **how to leverage big data volumes**. Datasets have grown exponentially the past few years and this is causing huge challenges for both the technologies as for the quantitative methods used. Statistical significance in large datasets gets a whole new dimension. In the traditional OR/MS approach, the pressing concern was often to collect enough data to find significant effects. The opposite is true for large datasets: almost every relationship can be measured as significant at the 5 percent level (Mortenson, Doherty, & Robinson, 2015). This urges to rethink established standards and reassess what methods can be used for hypothesis testing and model validation in such big datasets. Another question rises in the contradiction of the simplicity of an OR/MS model and the complexity and variety within a big dataset and how these can be combined to get value of this combination.

A second topic is the **utilization of new data architectures**. Many new types of databases, architectures and techniques arise and are at the technological end of the spectrum, whereas OR/MS applications are rather standard and straightforward. The need to demonstrate how these applications can be aligned within this new architectural framework in scientific literature is called upon by many authors (Anagnostou & Taylor, 2017; Mortenson, Doherty, & Robinson, 2015). Examples of Big Data applications within distributed systems are numerous, but case studies and reports of experimentation that explore the opportunities of OR/MS applications integration are recommended for future research.

Another managerial issue that should be addressed in future research is the **incorporation of unstructured data**. Traditional OR/MS models handle almost exclusively structured data, but they could benefit widely from the additional value of unstructured data (Gorman & Klimberg, 2014). The three main questions that need to be addressed are how such data is pre-processed, how such data can be used effectively in OR/MS models without adding too much complexity and which case studies can demonstrate the use of such data in OR/MS applications (Mortenson, Doherty, & Robinson, 2015).

A fourth research objective that should be an urging topic on the research agenda is that of **streaming data and real-time analytics**. This is already discussed in scientific literature to a certain extent, but there are still some unaddressed issues regarding the opportunities of modeling and data collection/processing to occur in real-time (Hazen, Skipper, Boone, & Hill, 2016). The autonomous flow of data into the OR/MS models should also be further explored. Case studies and literature reviews on these topics are in high demand[46].

Another issue that has not been addressed in scientific literature and that is a major challenge for the incorporation of Big Data Analytics into OR/MS models is the **visualization of data**. This topic has been around for quite some time, but is recently becoming an area of significant growth along the rise of Big Data. OR/MS can find genuine benefits from Big Data visualization, as this discipline is very much focused on decision making (McCauley & Graves, 2017). Examples of these benefits are the facilitation of model

[46] http://www.comsoc.org/tnsm/cfp/si-bdm

validation or increased transparency and acknowledgement from stakeholders. The transition from traditional result visualization towards the visualization in the Analytics age has not been delineated widely in literature and could benefit from further research. The combination of Big Data Analytics with virtual and augmented reality (see *section 6.3 Opportunities*) could also level up data visualization and could facilitate the closer relationship between Big Data Analytics and OR/MS.

The final topic that can benefit from further research is **OR/MS within the wider ecosystem.** Seen from a business perspective, it is in combination with other fields such as Big Data and Analytics that OR/MS can have greater impact and influence in the future. The encouragement of future collaborative research between these disciplines could be mutually beneficial for the wider ecosystem and the effectiveness and impact of the OR/MS methodology (Mortenson, Doherty, & Robinson, 2015).

6.2 Future Challenges

Many challenges inherent to Big Data and Analytics are already discussed for each business domain in *Chapter 5*. However, there are some general challenges that will stay and even some challenges that will pop up in the future. These are already subcutaneous present, but for the time being, they are not as crucial as they will be. First of all, privacy, security and ethics are discussed. These are interrelated because they are related to the people aspect of Big Data Analytics. Privacy issues and violations occurred more and more in the past few years, along with the trend towards more digitalization. Closely related to this issue is the security challenge. How can companies secure their data, including the data of their processes and customers, ensuring the privacy of the latter? Think of when big companies get hacked and customer data is stolen (email addresses, visa card data…), the worldwide cyberattack in May 2017[47], denial of service attacks... This is already a huge challenge at this moment in time, but will become even more urgent in the near future in Europe as a new General Data Protection Regulation (GDPR) will be in force starting from May 2018[48]. The ethics of Big Data are related to the previous two challenges and is addressed as well. After these challenges, technology and data characteristics challenges are explored. As Big Data Analytics is considered to be a competitive must-have, it is able to shape the technology landscape. IT infrastructures will take a more data-driven direction in the future. Another type of challenge addressed are the challenges related to the data characteristics of Big Data. The amounts of data will continue to grow, data will be produced faster than ever before and more and more variety will arise. Private and public companies, but governments as well will need to change the way of performing their daily business and the way they think about their processes if they want to keep up with the modern data-driven economy. Challenges inherent to OR are addressed last. The nature of OR/MS is so unique and simple, that the complexity of Big Data and Analytics causes many troubles.

Like any (relatively) new and promising field, Big Data (Analytics) must be viewed in terms of its capabilities but also in terms of its limitations with regards to the future. This section will try to give an overview of these limitations, but these are rather viewed as challenges, since Big Data Analytics is likely to evolve and cope with these challenges in the future. Most challenges are related to Big Data and Analytics, as these are penetrating the OR domain, causing corresponding challenges for OR/MS.

6.2.1 Privacy

As the collection of unstructured data becomes more economically viable, there is an incentive for organizations to collect as much data as possible. But it is not because people are willing to provide data

[47] http://www.bbc.com/news/world-europe-39907965
[48] http://www.eugdpr.org/

that this means that the use of Big Data is free from privacy implications (Nunan & Di Domenico, 2013). The massive collection of socioeconomic, behavioral, demographic, transactional and financial data for analytic purposes may lead to the decrease of civil liberties due to a loss of individual autonomy and privacy. The challenge thus is to ensure that people have sustainable control over their data, to prevent abuse and misuse by the people that collect or acquire these data, while preserving data utility.

For starters, it is very important to clarify that there is a huge difference concerning privacy regulations between the American and the European continent. In the United States, there is no clear privacy regulation and hence, companies and public services have more freedom on what to do with their data and how to acquire them. This is a substantial difference with the European privacy regulations. For the time being, there are still some different regulations between different European countries, but these differences are reducing in dribs and drabs. The reason behind this congruence is the European Community Law[49]. In 2016, the General Data Protection Regulation (GDPR) was adopted in Europe. Starting from May 2018, this Regulation will be fully applicable. It will be of crucial importance for private companies that do business in Europe to fine-tune their daily business and Big Data related applications with this new regulation in mind (Jacobs, 2017).

Ensuring personal data protection becomes more challenging as information is multiplied and shared even more dispersed around the world (European Commission, 2016b). Information ranging from a person's health, location, energy consumption, online activity and so forth can be publicized, raising concerns about profiling, discrimination, exclusion and loss of control. Of course, Big Data Analytics does not always involve personal data. But, when it does and this person lives in Europe, it should always comply with the rules and principles of data protection: the EU's Charter of Fundamental Rights says that everyone has the right to personal data protection in all aspects of life. Whether you are at work, at home, whilst shopping, when receiving medical treatment, at a police station or on the Internet, you should always be ensured that your data is protected and not just made publicly available.

Mayer-Schönberger & Cukier (2013) propose in their book four principles which could help to find a trade-off in this era of big (personal) data flow. These principles balance between the benefits and drawbacks of snooping around people's data:

- Privacy should be seen as a set of rules encompassing flows of information in ethical ways but not the ability to keep data secret
- Shared information can still be confidential

[49] http://ec.europa.eu/justice/data-protection/

- Big Data mining requires transparency
- Big Data can seriously threaten privacy

6.2.2 Security

Very closely related to the privacy challenge, is the security issue of Big Data Analytics. The biggest challenge for Big Data from a security point of view is the protection of the user's and data subject's privacy. Specifically, the issue around hacking or other forms of unauthorized access are considered a serious privacy abuse. Although computer systems, datahubs and Big Data-based infrastructures become more complex and better secured, computer-based systems are only as strong as their weakest point. The latter can be usually summarized in one word: humans. The regular access of several employees to all the company's data, how secured these may be for outsiders, is always subject to a breach to a certain extent. An excellent example is the WikiLeaks scandal. This scandal was caused by a low-level employee who copied data onto a fake Lady Gaga CD, while the US diplomatic network had one of the most advanced technical securities to protect this information (Leigh, 2010). Another important consideration in Big Data security are the problems concerning the cloud. Although clouds from giant providers are quite secure and safe, the fact remains that the data used by the company is no longer stored on the servers of the company itself but spread around the whole world. An organization then actually has no 100% guarantee on the security of their data. Especially governments will try to cope with this problem in the future. The Belgian governance for example is working on a "*G-cloud*"[50], a cloud for governmental data that will be spread only on servers that belong to governmental organizations in the European Union.

Up until now, data breaches we see in the press didn't have that much impact on people on a large scale although there has been an immense hack lately, called the '*WannaCry hack*'[51]. With the emergence of Big Data, breaches affecting these data can have more devastating consequences because this could potentially affect a much larger number of people (Lord, 2017). The consequences will not solely relate to a reputational point of view any more, but legal repercussions will become the standard. That is because more personal identifiable information (PII) is stored and can be concretized by the owners of the data (Marcello, 2016). This is a serious threat for communities and therefore, the challenge of data security will become more important than ever before. Lafuente (2015) identifies **five major considerations** when looking at Big Data security in the future: data anonymization, data encryption, access control and monitoring, data policy and finally governance frameworks.

[50] https://www.gcloud.belgium.be/nl/index.html
[51] https://en.wikipedia.org/wiki/WannaCry_ransomware_attack

The **anonymization of data** is important to make sure that privacy concerns are always addressed. It should be ensured that sensitive and personal information is removed from the set of collected records. The trade-off between data utility and privacy will be crucial and the right balance should be struck by companies. Before storing any data, they should be adequately anonymized such that unique identifiers of users are removed. Hence, it should be impossible that specific user identification is revealed during the analysis phase. There is however a huge footnote to this approach: anonymized data could be cross-referenced with other available data (which happens very often in Big Data Analytics) and then there is no guarantee that these will remain anonymous (Lafuente, 2015). Therefore, data encryption is paramount.

Data encryption is the main solution to ensure that data remains protected. This sounds an easy solution, but this is not the case at all. Organizations that have set up their infrastructure in the cloud need to perform their Big Data operations in this cloud environment and here, data cannot be sent encrypted by the users. This problem can be solved by Fully Homomorphic Encryption (FHE), which can ensure that the cloud can perform the Big Data operations without knowledge of the underlying plain text data, but this adds an extra complexity for Big Data scientists to set up this encryption. Therefore, this is often neglected in private companies as this costs time and money (Zibouh, Dalli, & Drissi, 2016).

The third major consideration is the **access of control to and monitoring of the data**. An access control policy can protect the information available in an organization by only allowing decryption of encrypted information if the entity trying to access it is authorized (Wu, Zhu, Wu, & Ding, 2014). At this moment in time, Big Data solutions and software offered by giant industry leaders do not always come with user authentication by default. Next to this, real-time security monitoring is key to make sure that no unauthorized access to the data is carried out. A threat intelligence system should also ensure that more sophisticated 'attacks' are detected and that the organization is capable to react to threats such as malware, vulnerabilities and bugs (Nair & Puri, 2015).

Another important consideration is **policies and compliance**. There are vendors that even start to offer compliance toolkits designed to work in a Big Data environment and to help organizations complying to the prevailing standards and rules. Risk assessments should be run now and then, privacy protection policies should be established and conclusive contracts should be drawn up when teaming up with other organizations and when data is shared (Cuquet, Vega-Gorgojo, Lammerant, Finn, & Hassan, 2017).

A final important consideration in the Big Data Analytics security landscape is that of **governance and frameworks**. If an adequate governance framework is not applied to Big Data applications or processes, then the data collected could be misleading and can even cause unexpected costs (Lafuente, 2015). The main problem to establish an adequate Big Data governance framework is that the concept of Big Data Analytics is relatively new and thus, no straightforward procedures or policies have ever been published. To conquer this challenge, a trial-and-error process will probably need to take place.

6.2.3 Ethics

In order to get a grasp on the ethics of Big Data, a clear theoretical background of Big Data Analytics is needed. This can be found in *section 2.1 Big Data Analytics*. There are a few so-called moral grey areas in Big Data Analytics, where complex negotiations about the relationship between ethics and epistemology are shaping the adoption of Big Data. It is often implicitly assumed in a Big Data project that data stays put within a specific context and temporal timeframe. In research with subjects, an informed consent is the standard point of ethical exchange between researchers and their subjects. This is signed before the research takes place. However, data creation is a process that is extended in time and across spatial and institutional setting, according to Helles & Jensen (2013). As it becomes cheaper to collect, store, analyze and re-analyze Big Datasets, it has become clear that informed consents at the beginning of the research cannot adequately capture the possible benefits and risks of consenting to the uses of a subject's data (Metcalf, Keller, & Boyd, 2016). Thus, it often becomes an ethical issue and this rises a debate on how to handle consent in Big Data research and public Big Data usage. The ethicality of Big Data is an urgent challenge now, but will remain a huge challenge in the future, as Big Data evolves rapidly.

The ethics of using Big Data and making Big Data analyses is not limited to the analysts or data collectors. Ethical behavior is a broader, organizational issue that needs to be addressed at all levels in an organization. The problem is that in many industries, this viewed as just an IT problem. But Big Data will stretch to the entire company in the future, as more and more companies will be fully data-integrated in a couple of years. This shows the importance of already addressing this challenge now. If there is not an ethical mindset present with the people that collect, process and analyze these data, how could it be present in an entire organization? The board, C-level executives, managers at all levels and by extension all employees should have an ethical compass when performing their daily business. Ethicality is an issue of all time, but in a digitalized age with data floating around everywhere, it becomes more important to make a distinction between morally correct and incorrect. Often, there is no awareness that something isn't ethically correct. More and more however, ethical guidelines and certifications pop up with the purpose of raising awareness in organizations. It is believed that companies have a better chance of promoting ethical behavior by their employees if they define what ethical behavior is. In a fast-changing digital environment, this will be a hard challenge. But hard challenges are not impossible of course.

6.2.4 Technology

Big Data Analytics has gradually become a critical top-line business issue that enterprises must tackle in order to remain competitive and relevant (Bantleman, 2013). Many organizations have changed their way of solving problems towards a more data-driven and more informed approach, which needs a change in most of the established IT systems at place. Big Data Analytics shapes IT infrastructure and this will be even more the case in the future. Most organizations do not yet have the appropriate skills sets or systems to take full advantage of the data they collect. Companies like Dell, IBM, Intel and Cisco Systems all see this as a huge opportunity to help their customers navigate in these unchartered water (Longoria, 2015). Established OR models will become more data-driven in the future, which will entail huge challenges to the existing technology.

The amount of data most organizations store, grows at a rate between 40 and 60 percent per year (Greengard, 2013). This growing rate of storage of course becomes a real challenge and becomes even more challenging each year, when looking to the future. Many organizations are looking at Big Data options like data lakes, which allow them to collect and store massive quantities of unstructured data in its native format (Crouch, 2016). The problem with data lakes is that they have to be constructed wisely or they quickly become a useless wasteland where data goes to never be retrieved again. Not many people have a lot of experience with these data lakes, which is thus a big challenge for the future.

Some estimates say there will be over 100 billion connected devices in 2020, with the amount of data collected doubling every two years (Howard, 2015). With this enormous amount of data, it is not realistic that IT infrastructure will stay the same for very long. The need for smarter and more flexible systems rises, as the evolution of Big Data Analytics will shape technology in the future (European Commission, 2016a). Applications, tools and infrastructure will all depend on how fast Big Data grows, how fast computers will need to compute and analyze and how important it is to keep up-to-date. This is also dependent on the data characteristics of Big Data, which brings us to the next challenge.

6.2.5 Data Characteristics

The characteristics of Big Data are determined by the 3V-model (*section 2.1 Big Data Analytics*). Volume, velocity and variety are inherent to Big Data and are the main differentiator between ordinary data and Big Data or between data processing and Big Data Analytics. As already mentioned several times, data will continue to grow at a high pace. The volume of available data for private companies and government poses many challenges as we speak. However, this will remain a challenge in the future because companies and governments will take measures to cope with this challenge but they will continue to lag behind the fast evolution of Big Data. At some point in time, like all technological evolutions, this will stagnate but when this will be, is still very unpredictable. Closely related to these data characteristic challenges is the quality of the data. In order to create real business value, a qualitative (Big Data) analysis needs to be executed to take informed decisions. The quality of this analysis depends entirely on the richness of the data. As data amounts will continue to rise and the speed at which data is produced will continue to grow, it will be a challenge to keep the dataset rich and it will even be harder to decide which data to store and which data can be discarded. The fact that there will be various sources of data income and that these data will be in different formats, does not alleviate the challenge.

6.2.6 OR inherent Challenges

There are several challenges inherent to Operations Research that complicate the closer relationship between Big Data Analytics and OR. The first is the **lack of mutual understanding between researchers and managers**. This mutual understanding is required in order to implement the results of a study. Researchers however often do not fully understand the real-life problem and do not tackle it the right way. Managers from their side must understand the method that the researcher has applied at the conceptual level, in order to dissect, challenge and implement the result of the study (Monks, 2016). Users of the designed model need to understand how the model works and why it is producing certain results. All these models need an appropriate amount of data before they can be validated, so a clear understanding of this data is also in place.

A second challenge is to **make policies and decisions an integral part of analysis**, not an afterthought. This changes the flow of daily business that has been around for quite some time. OR is unique in its origins and content, as it derives almost equally from deduction and induction (Sen, et al., 2014). OR has the capability to integrate with fields such as Big Data and Analytics and helps build scalable models and methods for many real-world applications. This integration will not go as smoothly as hoped for many organizations and endures a certain time and learning phase.

The third major challenge inherent to OR that impedes a closer relationship between Big Data Analytics and OR is the **nature of OR models**. These are always rather simplistic and deductive, whereas Big Data assumes a large and complex dataset, linked to an inferential model. The variability of data can cause problems in traditional OR models and therefore, the golden mean lies in between these two models. This implies a disruptive change in the whole OR approach, which has been established for over decades. This brings along other data-related challenges and there will be a lot of reluctance of modelers and researchers to change their usual way of working.

We can conclude that there are a lot of challenges lying ahead of using Big Data and Analytics in an OR/MS environment in the future. Companies and governments will have to adapt accordingly to address these challenges and to capture the value that Big Data Analytics can entail. A lot of opportunities lie in the offing of Big Data Analytics, ready to be taken advantage of. This brings us to the final section of this dissertation's body, namely the opportunities of Big Data and Big Data Analytics in an OR environment.

6.3 Opportunities

Before diving in the opportunities for Big Data Analytics in OR in the future, we will first take a look at some predictions and estimates made by some industry giants regarding Big Data Analytics, summarized in a Forbes article by Columbus (2016):

- Big Data and business analytics software revenues worldwide will grow from $122 billion in 2015 to more than $187 billion in 2019 (IDC, 2016)
- The total data market is expected to almost double in size, from $70 billion in revenue in 2015 to $132 billion in 2020 (Zwakman, Aslett, Stamper, Curtis, & Roy, 2016)
- By 2020, it is expected that predictive and prescriptive analytics will attract 40% of organizations' net new investment in business intelligence (Laney & Jain, 2016)

Given these estimates and predictions, it is clear that the industry and research see a bright future for Big Data and Analytics. This of course creates opportunities at multiple perspectives. Many opportunities will emerge when businesses will further explore areas such as **Artificial Intelligence (AI), Virtual Reality (VR) and the Internet of Things (IoT)** along the evolution of Big Data and Analytics. These areas can create synergies for companies and governments whereof the many different opportunities cannot be foreseen. Opportunities for Big Data and Analytics logically create opportunities for OR/MS as well, since these tend to intertwine more often than not.

The convergence of Big Data Analytics with **artificial intelligence** is inevitable as the automation of smarter decision-making is one of the next evolutions of Big Data (Canton, 2016). AI can be loosely defined as the science of making computers do things autonomously that require intelligence when done by humans. They are machines or computers that we consider 'smart'. An important application of AI is Machine Learning, which grants machines access to data to let them learn themselves (see *section 5.2.1 Applications in Healthcare – Techniques*). Considering the characteristics of Big Data, it is impossible to keep processing, analyzing and extracting knowledge at the same rate data grows. Therefore, AI and machine learning are needed, such that smart computers take on these tasks. The development and adoption of this technology is still in its early phase, such that there are many issues and these systems need a lot of human management for now, but this is likely to change in the near future (Helbing, et al., 2017). Decision support and increasingly autonomous decision making already is the norm, but more and more this will be done by machines and smart computer systems. The opportunities this brings along are difficult to quantify, but they will be big. *Table 25* gives a summary of recent scientific literature as regards general literature, future perspective and applications of artificial intelligence (often in combination with Big Data).

Literature	(Barbosa & Milios, 2015), (Bostrom, 2014), (Domingos, 2015), (Fermé & Leite, 2014), (Flasinski, 2016), (Kobbacy, Vadera, & Rasmy, 2007), (Lee, 2015), (Müller, 2016), (Sigaud & Buffet, 2013), (Torra & Narukawa, 2015), (Yang, 2013)
Future Perspective	(Armstrong, Sotala, & Ó hÉigeartaigh, 2016), (Atkinson, 2016), (Baciu, Opre, & Riley, 2016), (Bundy, 2017), (Chandra, 2016), (Goertzel, 2014), (Michael, 2015), (Müller & Bostrom, 2016)
Applications	International Journal of Engineering Applications of Artificial Intelligence, (Laalaoui & Bouguila, 2015), (Liang, et al., 2016), (Najafabadi, et al., 2015), (Renzi, Leali, Cavazzuti, & Andrisano, 2014), (Wauters & Vanhoucke, 2017), (Zang, Zhang, Di, & Zhu, 2015)

Table 25: Literature Overview Artificial Intelligence

Despite the technological development of the modern world, human interaction, judgement and logical reasoning are necessary while working with Big Data and performing analytics tasks. Therefore, human perceptional limitations are present during the analysis, limiting the opportunities of Big Data Analytics in an OR/MS environment. Big Data visualization can help overcome these limitations by visualizing the data. This is nothing new, as this is already done today on a daily basis. However, it is still rather limited as regards different types of visualizations and dimensions. The human brain has difficulties to capture this visualization and to draw logical conclusions out of it. This is where **virtual and augmented reality (VR/AR)** come into play. The capabilities of both technologies bring huge opportunities when combined with Big Data and Analytics, resulting in Big Data visualization. From an operations perspective, it will become far simpler to model the results and to visualize them. There is a slight difference between virtual and augmented reality: virtual reality is an artificial, computer-generated simulation or recreation of a real-life environment or situation, whereas augmented reality is a technology that layers computer-generated enhancements atop an existing reality in order to make it more meaningful through the ability to interact with it. The difference is that virtual reality offers a digital recreation of a real-life setting, while augmented reality delivers virtual elements as an overlay to the real world (Lindsay, 2015). A clear visual representation of a Big Data analysis is crucial for its interpretation and thus key for getting to the right conclusion and to take the right actions. These VR/AR applications are designed such that they understand the issues related to human perception and limited cognition. In the paper of Olshannikova, Ometov, Koucheryavy, & Olsson (2015), the different visualization tools are evaluated and it is outlined how they can improve Big Data visualization. These tools should facilitate visualization and the comprehension of the data analysis. The opportunity this brings along is that far more complex and more difficult problems can be analyzed, while managers and policy makers will still be able to understand the whole problem and analysis. Some applications that already exist today are summarized in the article of Boulos, Lu, Guerrero, Jennett, & Steed (2017). *Table 26* gives an overview of recent scientific literature concerning the general literature, future perspective and applications of VR and AR (often combined with Big Data and Analytics).

Literature	(Berg & Vance, 2017), (Garciá-Hernández, Anthes, Wiedemann, & Kranzlmüller, 2016), (Grandi, 2017), (Olshannikova, Ometov, Koucheryavy, & Olsson, 2015), (Schmalstieg & Höllerer, 2017), (Tan & Kim, 2016), (Zhaparov & Nassen, 2016)
Future Perspective	(Baciu, Opre, & Riley, 2016), (Cheng & Tsai, 2013), (Chi, Kang, & Wang, 2013), (Grajewski, et al., 2015), (Gutiérrez-Maldonado, Wiederhold, & Riva, 2016), (Lamberti, et al., 2014), (Lawson, Salanitri, & Waterfield, 2016), (Wang, Kim, Love, & Kang, 2013)
Applications	(Chi, Kang, & Wang, 2013), (Grajewski, Górski, Zawadzki, & Hamrol, 2013), (Lizcano, Manchado, Gomez-Jauregui, & Otero, 2017), (Novotny, Lacko, & Samuelcik, 2013), (Ong & Nee, 2013), (Sampaio & Martins, 2014), (Shumaker & Lackey, 2014)

Table 26: Literature Overview Virtual & Augmented Reality

Thirdly**, Internet of Things** refers to day-to-day appliances like air purifiers, cameras or light bulbs that are connected to the internet for remote management. They can even be programmed to turn on when the air for example contains contaminants or when a motion is detected (Tham, 2017). They are smart solutions, connected to the internet that produce many data and can make autonomous decisions, based on that data. Those 'things' can communicate with each other and interchange data. The combination of Big Data Analytics with IoT is undoubtedly one of the biggest opportunities to grasp in the next decade. Big Data Analytics and IoT are converging, as they are combined more often than not. The intersection of IoT with geospatial Big Data (data with explicit geographic positioning information included within) lies in the reality of sensors on the ground, coupled with near real-time modeling of visible spectrum data gathered from remote sensing. More and more, intelligent Big Data applications will connect with different databases at the same time, facilitating the emergence of IoT. According to Dasgupta (2017), it is about "*connecting the entire intelligent things and then making location sense out of it*". *Table 27* gives a concise summary of recent scientific literature on IoT, categorized in general literature, future perspective and applications of IoT.

Literature	(Alam, Katsikas, Beltramello, & Hadjiefthymiades, 2017), (Chandrakanth, Venkatesh, Mahesh, & Naganjaneyulu, 2014), IEEE Internet of Things Journal, (O'Donovan, Leahy, Bruton, & O'Sullivan, 2015), (Ray, 2017), (Stokjoska & Trivodaliev, 2017), (Venkatesh, 2017), (Wortmann & Flüchter, 2015), (Xu, He, & Li, 2014)
Future Perspective	(Gubbi, Buyya, Marusic, & Palaniswami, 2013), (Jin, Gubbi, Marusic, & Palaniswami, 2014), (Sicari, Rizzardi, Grieco, & Coen-Porisini, 2015), (Singh, Tripathi, & Jara, 2014), (Stankovic, 2014), (Whitmore, Agarwal, & Xu, 2015)
Applications	(Al-Fuqaha, Guizani, Mohammadi, Aledhari, & Ayyash, 2015), (Girau, Martis, & Atzori, 2017), (Hsieh, Chang, Wang, Chen, & Chao, 2016), (Kamienski, et al., 2017), (Lin, et al., 2017), (Maarala, Su, & Riekki, 2017), (Meng, Wu, Muvianto, & Gray, 2017), (Rymaszewska, Helo, & Gunasekaran, 2017), (Suthakar, Magnoni, Smith, Khan, & Andreeva, 2016)

Table 27: Literature Overview Internet of Things

The main angle of incidence in this opportunities section is that of Big Data Analytics. This is because we are convinced that the opportunities of Big Data and Analytics are much more varied than the opportunities of OR. The opportunities that will come up for Big Data will become opportunities for OR as well, via the closer relationship between the two.

7. Discussion

7.1 Summary

Data is flooding our daily lives and will continue to do so in the future. Numerous applications already exist nowadays that are built on operational models and whereon exuberant amounts of data is run. This creates unseen **opportunities** in all business areas imaginable, but also brings dubious **challenges**. The answers to the following not so unimportant questions will be vital for the future direction of Big Data Analytics and its usage in Operations Research: “What will happen with privacy regulations and governance?”, “Will organizations be able to secure their data and information from outsiders?” and “How ethically correct will the future of Big Data Analytics be in an OR environment?”. The opinion of the big mass on these subjects is essential and will demarcate future opportunities.

Big Data Analytics in **marketing** models is already reasonably entrenched. Applications can be divided into categories such as sentiment analytics, customer 360, customer segmentation, next best offer and channel journey. The main goal in marketing applications is to get to know the customers better and to approach them accordingly. Organizations learn from their customers and adjust to the market, which benefits society in general, while it also leverages organizational impact and value.

In the **healthcare** area, modeling and research has benefited to a certain extent from Big Data Analytics recently, but still has a lot more to offer. The greatest value can be revealed by further improving the incorporation of Big Data and Analytics in the daily management of hospitals and healthcare institutions. Applications that exist today can be categorized in research, preventive analytics, fraud detection and planning and scheduling. The future of healthcare will be even far more data-driven than it is now, which will result in immeasurable chances and groundbreaking discoveries.

Operations and supply chain management already takes advantage of predictive maintenance, process improvement and risk management by incorporating Big Data Analytics into their operational pursuits. This domain could benefit widely by further exploring the integration of machine learning, artificial intelligence and internet of things into their established systems.

The **public sector** is a bit different from the other discussed business domains, as these organizations are all public instead of private. This implies additional challenges and higher transparency. The most important applications or areas of applications that are prominent today are smart cities, public data hubs and security intelligence. The public sector disposes of the most information and is actually sitting on a sleeping gold mine of data, that is waiting to get excavated. In order to do this, the public sector still needs to evolve a lot towards a more data-driven and information-based decision making process.

A lot has been written on Big Data, Analytics and Operations Research, but there is a lot more to investigate and that can be subject of **future research**. The combination of the above domains is by far not fully exploited in literature and in real-life and therefore, the payoff of certain applications is up until today rather marginally incremental. More value can be generated by further exploring the opportunities of Big Data Analytics in an OR environment.

7.2 Research Contribution

The contribution of this dissertation to scientific research is threefold. The main contribution of this literature study is the outline on how the adoption of Big Data Analytics in the field of Operations Research is evolving, based on what has already been written. Important things that need to be considered when starting to adopt Big Data and Analytics in an organization are addressed. Hence, the relationship between the two formal separate domains becomes more coherent and explicit. This dissertation also makes interested readers and people in the corresponding industries aware of the challenges and trends that are merging in their field or in general. A third important contribution is the overview of Big Data Analytics applications given in four important business domains. This could contribute to an overview of all applications, which could be the subject of a doctoral thesis or extensive research.

The unique feature of this dissertation is that it not only explores the relevant scientific and technical literature, but the relevant managerial literature as well and even opinions of prominent people in a particular field, expressed in blogs. The value of this dissertations propagates from the fact that the overview of applications, opportunities and trends is not limited to just one business field.

7.3 Limitations

A limitation of this dissertation is that not all business domains are assessed. As already mentioned, the applications of the assessed business domains can be seen as the epitome of other business domains. A more extensive and comprehensive overview could be the subject of a larger research project or book, whereat this research could contribute.

Another limitation is that not all relevant articles are gathered. This is an almost impossible task as the literature on Big Data, Analytics and Operations Research is so widespread, extensive and updated daily. The literature that links Big Data Analytics with Operations Research is only a very small part of this literature, such that further research on the relationship between the two could have had a positive impact on this research as well.

Reference List

Adomavicius, A. (2014, December 19). *Using Big Data from the IoT to predict machine failure.* Retrieved April 14, 2017, from http://www.manufacturing.net/article/2014/12/using-big-data-iot-predict-machine-failure

Agrawal, A. (2016, February 19). *How Marketing Has Changed and Why It Matters.* Retrieved April 3, 2017, from Inc.: http://www.inc.com/aj-agrawal/how-marketing-has-changed-and-why-it-matters.html

Alam, M. F., Katsikas, S., Beltramello, O., & Hadjiefthymiades, S. (2017). Augmented and virtual reality based monitoring and safety system: A prototype IoT platform. *Journal of Network and Computer Applications*, 1-11. doi:10.1016/j.jnca.2017.03.022

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology, 22(1)*, 3-21. doi:10.1080/10630732.2014.942092

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communication Surveys & Tuturials, 17(4)*, 2347-2376. doi:10.1109/comst.2015.2444095

Anagnostou, A., & Taylor, S. J. (2017). A distributed simulation methodological framework for OR/MS applications. *Simulation Modelling Practice and Theory, 70*, 101-119. doi:10.1016/j.simpat.2016.10.007

Armstrong, S., Sotala, K., & Ó hÉigeartaigh, S. (2016). Errors, Insights, and Lessons of Famous Artificial Intelligence Predictions - And What They Mean for the Future. In V. C. Müller , *Risks of Artificial Intelligence* (pp. 29-68). Chapman and Hall/CRC.

Arnott, D., & Pervan, G. (2008). Eight key issues for the decision support systems discipline. *Decision Support Systems, 44(3)*, 657-672. doi:10.1016/j.dss.2007.09.003

Arnott, D., & Pervan, G. (2014). A critical analysis of decision support systems research revisited: the rise of design science. In L. P. Willcocks, C. Sauer, & M. C. Lacity, *Enacting Research Methods in Information Systems - Volume 3* (pp. 43-103). Springer International Publishing. doi:10.1007/978-3-319-29272-4

Assunçao, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., & Buyya, R. (2015). Big Data Computing and Clouds: Trends and Future Directions. *Journal of Parallel and Distributed Computing, Vol 79-80*, 3-15. doi:10.1016/j.jpdc.2014.08.003

Atalla, G., Banks, B., Littlejohn, M., & Hiscock-Croft, R. (2016). *The Power of Three for smarter, more resilient cities.* EYGM Limited.

Atkinson, R. D. (2016). *"It's going to kill us!" and Other Myths About the Future of Artificial Intelligence.* Information Technology & Innovation Foundation.

Attendance. (2015). *Optimize Employee Scheduling with Time and Attendance.* Retrieved April 13, 2017

Auschitzky, E., Hammer, M., & Rajagopaul, A. (2014, July). *How big data can improve manufacturing.* Retrieved April 14, 2017, from McKinsey & Company: http://www.mckinsey.com/business-functions/operations/our-insights/how-big-data-can-improve-manufacturing

Baciu, C., Opre, D., & Riley, S. (2016). A New Way of Thinking in the Era of Virtual Reality and Artificial Intelligence. *Educatia 21, 14*, 43-51.

Bantleman, J. (2013, January). *Big Data: Business or Technology challenge.* Retrieved April 29, 2017, from Wired: https://www.wired.com/insights/2013/01/big-data-business-or-technology-challenge/

Barbosa, D., & Milios, E. (2015). Advances in Artificial Intelligence. *28th Canadian Conference on Artificial Intelligence* (pp. 69-138). Halifax, Nova Scotia, Canada: Springer International Publishing.

Barceló, J. (2015). Analytics and the art of modeling. *International Transactions in Operational Research, 22(3)*, 429-471. doi:10.1111/itor.12165

Barclay, A. (n.d.). *Using Hadoop to detect health care fraud, waste and abuse.* Retrieved April 12, 2017, from MapR: https://mapr.com/customers/unitedhealthcare/

Barton, D., & Court, D. (2012). Making Advanced Analytics Work For You. *Harvard Business Review, 90(10)*, 78-83.

Basu, A. (2013). Five pillars of prescriptive analytics success. *Analytics Magazine*, 8-12.

Batarseh, F. A., & Latif, E. A. (2015). Assessing the Quality of Service Using Big Data Analytics With Application to Healthcare. *Big Data Research*. doi:10.1016/j.bdr.2015.10.001

Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. *Information Fusion, 28*, 45-59. doi:10.1016/j.inffus.2015.08.005

Berardi, U. (2013a). Clarifying the New Interpretations of the Concept of Sustainable Building. *Sustainable Cities and Society, 8*, 72-78. doi:10.1016/j.scs.2013.01.008

Berardi, U. (2013b). Sustainability Assessments of urban Communities through Rating Systems. *Environment, Development and Sustainability, 15(6)*, 1573-1591. doi:10.1007/s10668-013-9462-0

Berg, L. P., & Vance, J. M. (2017). Industry use of virtual reality in product design and manufacturing: a survey. *Virtual Reality, 21(1)*, 1-17. doi:10.1007/s10055-016-0293-9

Berg, N. (2014, June 25). *Predicting crime, LAPD-style.* Retrieved April 22, 2017, from The Guardian: https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report

Berry, M. J., & Linoff, G. S. (2011). Chapter 16 - Link Analysis. In M. J. Berry, & G. S. Linoff, *Data mining techniques: for marketing, sales, and customer relationship management (3th edition).* John Wiley & Sons.

Bertolucci, J. (2013, March 18). *Intel cuts manufacturing costs with Big Data.* Retrieved April 14, 2017, from InformationWeek: http://www.informationweek.com/software/information-management/intel-cuts-manufacturing-costs-with-big-data/d/d-id/1109111

Beuder, J. (2013, November 5). *How can big data improve the customer experience this holiday season?* Retrieved April 8, 2017, from ICMI: http://www.icmi.com/Resources/Customer-Experience/2013/11/How-Can-Big-Data-Improve-The-Customer-Experience-This-Holiday-Season

Bort, J. (2016, December 7). *How IBM Watson saved the life of a woman dying from cancer, exec says.* Retrieved April 13, 2017, from Business Insider UK: http://uk.businessinsider.com/how-ibm-watson-helped-cure-a-womans-cancer-2016-12?r=US&IR=T

Bose, R. (2009). Advanced analytics: opportunities and challenges. *Industrial Management & Data Systems, 109(2)*, 155-172. doi:10.1108/02635570910930073

Bostrom, N. (2014). *Superintelligence.* Oxford, UK: Oxford University Press .

Boulos, M. N., Lu, Z., Guerrero, P., Jennett, C., & Steed, A. (2017). From urban planning and emergency training to Pokémon Go: applications of virtual reality GIS (VRGIS) and augmented reality GIS (ARGIS) in personal, public and environmental health. *International Journal of Health Geographics, 16(1)*, 1-11. doi:10.1186/s12942-017-0081-0

Brahm, C., Cheris, A., & Sherer, L. (2016, August 8). *What Big Data Means for Customer Loyalty.* Retrieved May 9, 2017, from Bain & Company: http://www.bain.com/publications/articles/what-big-data-means-for-customer-loyalty.aspx

Brennan, N., Oelschlaeger, A., Cox, C., & Tavenner, M. (2014). Leveraging the Big Data revolution: CMS is expanding capabilities to spur health system transformation. *Health Affairs, 33(7)*, 1195-1202. doi:10.1377/hlthaff.2014.0130

Brown, D. E., Famili, F., Paass, G., Smith-Miles, K., Thomas, L. C., Weber, R., . . . Maldonado, S. (2011). Future trends in business analytics and optimization. *Intelligent Data Analysis, 15(6)*, 1001-1017. doi:10.3233/IDA-2011-0506

Bughin, J. (2016). Big data, Big bang? *Journal of Big Data, 3(1)*, 1-14. doi:10.1186/s40537-015-0014-3

Bundy, A. (2017). Preparing for the future of Artificial Intelligence. *AI & Society, 32(2)*, 285-287. doi:10.1007/s00146-016-0685-0

Cameron, N. (2015, September 14). *Jeanswest details how it's building personalised digital engagement.* Retrieved April 8, 2017, from CMO: http://www.cmo.com.au/article/584323/jeanswest-details-how-it-building-personalised-digital-engagement/

Canton, J. (2016, July 5). *From Big Data to Artificial Intelligence: The next digital disruption.* Retrieved May 1, 2017, from Huffington Post: http://www.huffingtonpost.com/james-canton/from-big-data-to-artifici_b_10817892.html

CBPL: Commissie voor de bescherming van de persoonlijke levenssfeer. (2016). *Big Data Rapport AH-2016-0154.*

Celdrán Bernabeu, M., Mazón López, J., Giner Sánchez, D., & Ivars Baidal, J. (2016). Big Data and Smart Tourism Destinations: Challenges and opportunities from an industry perspective. *School of Hospitality and Tourism Management Conference* (pp. 1-16). Surrey (UK): University of Surrey.

Celebi, E. M., Kingravi, H. A., & Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications, 40*, 200-210. doi:10.1016/j.eswa.2012.07.021

Chan, S. H., Song, Q., Sarker, S., & Plumlee, D. R. (2017). Decision support system (DSS) use and decision performance: DSS motivation and its antecedents. *Information & Management*. doi:10.1016/j.im.2017.01.006

Chand, S. (n.d.). *Top 6 steps involved in Operations Research - Explained.* Retrieved March 20, 2017, from Your Article Library: http://www.yourarticlelibrary.com/ergonomics/operation-research/top-6-steps-involved-in-operation-research-explained/34688/

Chandra, M. (2016). Artificial Intelligence and the Future of Knowledge Workers. *5th International Conference on Reliability, Infocom Technologies and Optimization* (p. 44). Uttar Pradesh, Noida, India: IEEE.

Chandrakanth, S., Venkatesh, K., Mahesh, J. U., & Naganjaneyulu, K. V. (2014). Internet of Things. *International Journal of Innovations & Advancement in Computer Science, 3(8)*, 16-20.

Chawla, N. V., & Davis, D. A. (2013). Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework. *Journal of General Internal Medicine, 28(3)*, 660-665. doi:10.1007/s11606-013-2455-8

Chen, C. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences, 275*, 314-347. doi:10.1016/j.ins.2014.01.015

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: from Big Data to Big Impact. *MIS Quarterly, 36(4)*, 1165-1188.

Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications, 19(2)*, 171-209. doi:10.1007/s11036-013-0489-0

Chen, M., Mao, S., Zhang, Y., & Leung, V. C. (2014). *Big Data Related Technologies, Challenges and Future Prospects.* London, UK: Springer.

Cheng, K.-H., & Tsai, C.-C. (2013). Affordances of Augmented Reality in Science Learning: Suggestions for Future Research. *Journal of Science Education and Technology, 22(4)*, 449-462. doi:10.1007/s10956-012-9405-9

Chi, H.-L., Kang, S.-C., & Wang, X. (2013). Research trends and opportunities of augmented reality applications in architecture, engineering and construction. *Journal of Automation in Construction, 33*, 116-122. doi:10.1016/j.autcon.2012.12.017

Choi, T.-M., Chan, H. K., & Yue, X. (2017). Recent Development in Big Data Analytics for Business Operations and Risk Management. *IEEE Transactions on Cybernetics, Vol. 47 (1)*, 81-92.

Chopra, S., & Sodhi, M. S. (2014). Reducing the risk of supply chain disruptions. *MIT Sloan Management Review, 55(3)*.

Claessens, S., & Kodres, L. (2014). *The Regulatory Responses to the Global Financial Crisis: Some Uncofmortable Questions.* Research Department and Institute for Capacity Development.

Columbus, L. (2016, August 20). *Roundup Of Analytics, Big Data & BI Forecasts And Market Estimates, 2016.* Retrieved January 31, 2017, from Forbes Magazine: http://www.forbes.com/sites/louiscolumbus/2016/08/20/roundup-of-analytics-big-data-bi-forecasts-and-market-estimates-2016/#e4bba3949c5f

Coomer, S. (2016, February 22). *Amex Applications: The One (Targeted) Way to Avoid the "One Bonus Per Lifetime" Rule.* Retrieved April 8, 2017, from milestomemories: http://milestomemories.boardingarea.com/amex-credit-card-application-bonus-rules/

Corner, S. (2014, March 4). *Westpac using big data to woo customers with offers made to measure.* Retrieved April 8, 2017, from The Sydney Morning Herald: http://www.smh.com.au/it-pro/business-it/westpac-using-big-data-to-woo-customers-with-offers-made-to-measure-20140303-hvfx5.html

Crawford, K., Miltner, K., & Gray, M. L. (2014). Critiquing Big Data: Politics, Ethics, Epistemology. *International Journal of Communication, 8*, 1663-1672.

Crosman, P. (2013, July 15). *How Zions Bank is conquering big data for marketing campaigns.* Retrieved April 8, 2017, from American Banker: https://www.americanbanker.com/news/how-zions-bank-is-conquering-big-data-for-marketing-campaigns

Crouch, A. (2016, April 7). *The Top 5 Big Data Analytics Challenges Facing Big Business.* Retrieved May 1, 2017, from Aptera: http://blog.apterainc.com/business-intelligence/the-top-5-big-data-analytics-challenges-facing-big-business

Cuquet, M., Vega-Gorgojo, G., Lammerant, H., Finn, R., & Hassan, U. u. (2017). Societal impacts of big data: challenges and opportunities in Europe. *rXiv preprint arXiv:1704.03361*, 1-17.

Curry, D., Blijleven, W., & Van de Walle, S. (2014). *Current and future trends in public sector reform: the views of trade unions and consultants in ten European countries.* Cocops.

Dai, Y., & Sun, H. (2014). The naive Bayes text classification algorithm based on rough set in the cloud platform. *Journal of Chemical and Pharmaceutical Research, 6(7)*, 1636-1643.

Dasgupta, A. (2017, February 7). *The Continuum: Big Data, Cloud & Internet of Things.* Retrieved May 2, 2017, from IBM: https://www.ibm.com/blogs/internet-of-things/big-data-cloud-iot/

Datanami. (2013, February 4). *A holistic approach to big data analytics.* Retrieved April 13, 2017, from Datanami: https://www.datanami.com/2013/02/04/a_holistic_approach_to_big_data_analytics/

Davenport, T. H. (2006). Competing on Analytics . *Harvard Business Review*, 1-10.

Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning.* Cambridge: Harvard Business School Press.

Davenport, T. H., & Patil, D. J. (2012, October). *Data Scientist: The Sexiest Job of the 21st Century.* Retrieved April 10, 2017, from Harvard Business Review: https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at Work: Smarter Decisions, Better Results.* Cambridge : Harvard Business School Press.

Davis, T., Nayeri, M., Suryanarayan, S., & Wilson, G. (2015). *Trusting big data: Perspective on data governance as a customer analytics investment.* Deloitte Development LLC.

de Bruijn, A. (2013, December 19). *Naar een fraudebeeld Nederland - Inzicht in fraude draagt bij aan bewustwording en effectieve prioriteitsstelling in de aanpak.* Retrieved April 13, 2017, from PWC: https://www.pwc.nl/nl/assets/documents/pwc-naar-een-fraudebeeld-nederland.pdf

de Roys, S., Bouchard, A., Sehad, M., Pellegrinelli, R., Mantz, B., Xuereb, J.-M., & Viniane, C. (2017). *Hyper-personalization vs. segmentation: has big data made customer segmentation redundant.* Capgemini Consulting.

Deloitte. (2013). *Unleashing the power within Analytics-driven process design.* Deloitte Southeast Asia Ltd.

Demchenko, Y., Grosso, P., & de Laat, C. (2013). Addressing Big Data Issues in Scientific Data Infrastructure. *International Conference on Collaboration Technologies and Systems (CTS)* (pp. 48-55). IEEE. doi:10.1109/CTS.2013.6567203

Demchenko, Y., Turkmen, F., de Laat, C., Blanchet, C., & Loomis, C. (2016). Cloud based big data infrastructure: Architectural components and automated provisioning. *International Conference on High Performance Computing & Simulation (HPCS)* (pp. 628-636). IEEE. doi:10.1109/HPCSim.2016.7568394

Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems, 55(1)*, 412-421. doi:10.1016/j.dss.2012.05.048

Desouza, K. C. (2014). Realizing the Promise of Big Data - Implementing Big Data Projects. *Using Technology Series*, 23-29.

Dey, D., & Kumar, S. (2010). Reassessing data quality for information products. *Management Science, 56(12)*, 2316-2322. doi:10.1287/mnsc.1100.1261

Dhawan, R., Singh, K., & Tuteja, A. (2014, February). *When big data goes lean.* Retrieved April 15, 2017, from McKinsey & Company - Operations: http://www.mckinsey.com/business-functions/operations/our-insights/when-big-data-goes-lean

Dohr, A., Modre-Osprian, R., Drobics, M., Hayn, D., & Schreier, G. (2010). The Internet of Things for Ambient Assisted Living. *Seventh International conference on Information Technology: new generations (ITNG)* (pp. 804-809). IEEE.

Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our Wold.* New York, USA: Basic Books.

Drexler, M. (2014). Big data's big visionary. *Magazine of the Harvard T.H. Chan School of Public Health*. Retrieved April 11, 2017, from https://www.hsph.harvard.edu/magazine/magazine_article/big-datas-big-visionary/

Du, S., Hu, L., & Song, M. (2016). Production optimization considering environmental performance and preference in the cap-and-trade system. *Journal of Cleaner Production, 112*, 1600-1607. doi:10.1016/j.jclepro.2014.08.086

Dubey, R., & Gunasekaran, A. (2014). Agile manufacturing: framework and its empirical validation. *International Journal of Advanced Manufacturing Technology, 76(9-12)*, 2147-2157. doi:10.1007/s00170-014-6455-6

Dubois, L. (2015, July 1). *Data-influenced Marketing Decisions Weighing on Today's CMOs.* Retrieved April 3, 2017, from Trillium Insights: http://blogs.trilliumsoftware.com/trilliuminsights/2015/07/data-influenced-marketing-decisions-weighing-on-todays-cmos.html

Dutcher, R. (2014, March 6). *Linking Big Data to Big Process Improvement... An Imperative.* Retrieved April 14, 2017, from Capgemini: https://www.capgemini.com/blog/bpo-thought-process/2014/03/linking-big-data-to-big-process-improvementan-imperative

Ekinci, Y., Ülengin, F., Uray, N., & Ülengin, B. (2014). Analysis of customer lifetime value and marketing expenditure decisions through a Markovian-based model. *European Journal of Operational Research, 237(1)*, 278-288. doi:10.1016/j.ejor.2014.01.014

Elsevier. (2016). Retrieved March 1, 2017, from Special Issue on "Emerging Trends, Issues and Challenges in Internet of Things, Big Data and Cloud Computing": https://www.journals.elsevier.com/future-generation-computer-systems/call-for-papers/special-issue-on-emerging-trends-issues-and-challenges-in-in

Emani, C. K., Cullot, N., & Nicolle, C. (2015). Understandable Big Data: A survey. *Computer Science Review, 17*, 70-81. doi:10.1016/j.cosrev.2015.05.002

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research, 69(2)*, 897-904. doi:10.1016/j.jbusres.2015.07.001

European Commission. (2013). *Investing in health - Commission staff working document social investment package.* Brussels: European Commission.

European Commission. (2016a). *Communication from the commission to the European Parliament and the council - Stronger and Smarter Information Systems for Borders and Security.* Brussels: European Commission.

European Commission. (2016b). *The Eu Data Protection Reform and Big Data: Factsheet.* European Commission. Retrieved from http://ec.europa.eu/justice/data-protection/files/data-protection-big-data_factsheet_web_en.pdf

Experian. (2013). *Case Study: Vodafone integrates web, email and mobile to enhance prospective customers' online shopping experience.* Retrieved April 9, 2017, from http://www.experian.co.uk/assets/marketing-services/case-studies/case-study-vodafone.pdf

Farah, A. (2013). Big Data in Marketing. In L. Mawhinney, & R. Self, *Big Data for SMEs: Questions of Opportunities, Challenges, Benefits and Operations - 2nd edition* (pp. 12-15). Derby, UK: University of Derby.

Feloni, R. (2013, December 11). *The NSA is using Google's advertising cookies to track its targets.* Retrieved April 23, 2017, from Business Insider UK: http://www.businessinsider.com/the-nsa-uses-google-cookies-for-hacking-2013-12?IR=T

Fermé, E., & Leite, J. (2014). Logics in Artificial Intelligence. *14th European Conference on JELIA* (pp. 62-281). Funchal, Madeira, Portugal: Springer International Publishing.

Figueras, J., & Mayer, J. (2012, August 29). *Case Study: Eircom is investing in analytics to improve customer experience and drive growth.* Retrieved April 9, 2017, from https://www.scribd.com/document/205648749/Case-Study-Eircom-Improve-Customer-Experience

Finsterwalder, J. (2016). A 360-degree view of actor engagement in service co-creation. *Journal of Retailing and Consumer Services*, 1-3. doi:10.1016/j.jretconser.2016.08.005

Firestein, S. (2012). *Ignorance - How it drives science.* New York, USA: Oxford University Press.

Fita, M. (2016, June 16). *How Marketing Has Changed In the Past 20 Years.* Retrieved April 3, 2017, from brandignity: https://www.brandignity.com/2016/06/how-marketing-has-changed-in-the-past-20-years/

Flasinski, M. (2016). *Introduction to Artificial Intelligence.* Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-40022-8

Fogelman-Soulié, F., & Lu, W. (2016). Implementing Big Data Analytics Projects in Business. In N. Japhowicz, & J. Stefanowski, *Big Data Analysis: New Algorithms for a New Society* (pp. 141-158). Springer International Publishing Switzerland: Springer. doi:10.1007/978-3-319-26989-4_6

Forgione, G. A. (1990). *Quantitative Management.* Chicago: The Dryden Press.

Franková, P., Drahosova, M., & Balco, P. (2016). Agile project management approach and its use in big data management. *Procedia Computer Science, vol 83*, 576-583. doi:10.1016/j.procs.2016.04.272

Gahi, Y., Guennoun, M., & Mouftah, H. T. (2016). Big Data Analytics: Security and Privacy Challenges. *IEEE Symposium on Computers and Communication (ISCC)*. doi:10.1109/iscc.2016.7543859

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35(2)*, 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

Gantz, J., & Reinsel, D. (2012). The digital universe in 2020: Big Data, bigger digital shadows, and biggest growth in the far east. *IDC iVIEW*, 1-16.

Garciá-Hernández, R. J., Anthes, C., Wiedemann, M., & Kranzlmüller, D. (2016). Perspectives for Using Virtual Reality to Extend Visual Data Mining in Information Visualization. *Aerospace Conference* (pp. 1-11). Big Sky, MT, USA: IEEE.

Gartner. (2015, December 4). *The Chief Analytics Officer's Vision Sets the Narrative for the Business Analytics Strategy*. Retrieved May 2, 2016, from https://www.gartner.com/search/site/freecontent/simple?typeaheadTermType=&typeaheadTermId=&keywords=descriptive%2C+predictive+and+prescriptive+analytics

Gesundheit Österreich Forschungs- und Plannungs GmbH. (2016). *Study on Big Data in Public Health, Telemedine and Healthcare - Final Report.* Brussels: European Commission.

Girau, R., Martis, S., & Atzori, L. (2017). Lysis: A platform for IoT Distributed Applications Over Socially Connected Objects. *IEEE Internet of Things Journal, 4(1)*, 40-51. doi:10.1109/jiot.2016.2616022

Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence, 5(1)*, 1-46. doi:10.2478/jagi-2014-0001

Gorman, M. F., & Klimberg, R. K. (2014). Benchmarking Academic Programs in Business Analytics. *Interfaces, 44(3)*, 329-341. doi:10.1287/inte.2014.0739

Grajewski, D., Diakun, J., Wichniarek, R., Dostatni, E., Bun, P., Gorski, F., & Karwasz, A. (2015). Improving the skills and knowledge of future designers in the field of ecodesign using virtual reality technologies. *Procedia Computer Science, 75*, 348-358. doi: 10.1016/j.procs.2015.12.257

Grajewski, D., Górski, F., Zawadzki, P., & Hamrol, A. (2013). Application of Virtual Reality Techniques in Design of Ergonomic Manufacturing Workplaces. *Procedia Computer Science, 25*, 289-301. doi:10.1016/j.procs.2013.11.035

Grandi, J. G. (2017). Design of Collaborative 3D User Interfaces for Virtual and Augmented Reality. *Virtual Reality* (pp. 419-420). Los Angeles, CA, USA: IEEE.

Greengard, S. (2013, January 30). *Facing the Big Challenges of Big Data.* Retrieved Mary 1, 2017, from BizTech: http://www.biztechmagazine.com/article/2013/01/facing-big-challenges-big-data

Greenwald, G., & MacAskill, E. (2013, June 7). *NSA Prism program taps in to user data of Apple, Google and others.* Retrieved April 22, 2017, from The Guardian: https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

Greenwald, M. (2016, September 20). *Way beyond jeopardy: 5 marketing uses of IBM Watson.* Retrieved April 5, 2017, from Forbes Magazine: https://www.forbes.com/sites/michellegreenwald/2016/09/20/way-beyond-jeopardy-5-marketing-uses-of-ibm-watson/#216ad8db4a76

Groenfeldt, T. (2013, June 11). *Banks use big data to understand customers across channels.* Retrieved April 8, 2017, from Forbes Magazine: https://www.forbes.com/sites/tomgroenfeldt/2013/06/11/banks-use-big-data-to-understand-customers-across-channels/#898235e3218a

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. *Future Generation Computer Systems, 29(7)*, 1645-1660. doi:10.1016/j.future.2013.01.010

Gudipati, M., Rao, S., Mohan, N. D., & Gajja, N. K. (2013). Big Data: Testing approach to overcome quality challenges. *Infosys Labs Briefings, 11(1)*, 65-72.

Gudivada, V. N., Baeza-Yates, R., & Raghavan, V. V. (2015, March). Big Data: Promises and Problems. *Computer, 48(3)*, 20-23. doi:10.1109/MC.2015.62

Guerrieri, A. (2014, January 7). *Bank of China achieves customer-centric transformation with IBM.* Retrieved April 9, 2017, from IBM: http://www-03.ibm.com/press/us/en/pressrelease/42845.wss

Gutiérrez-Maldonado, J., Wiederhold, B. K., & Riva, G. (2016). Future directions: how virtual reality can further improve the assessment and treatment of eating disorders and obesity. *Cyberpsychology, Behavior, and Social Networking, 19(2)*, 148-153. doi:10.1089/cyber.2015.0412

Haag, S., Cummings, M., & Dawkins, J. (2000). *Management Information Systems for the Information Age.* McGraw-Hill Ryerson Limited.

Hall, R. (2013, April 14). *In big data we hope and distrust.* Retrieved February 19, 2017, from DSSResources: http://dssresources.com/papers/features/hall/hall04072013.htm

Harris, D. (2014, February 8). *3 lessons in big data from the Ford Motor Company.* Retrieved April 5, 2017, from https://gigaom.com/2014/02/08/3-lessons-in-big-data-from-the-ford-motor-company/

Harris, S. L., May, J. H., & Vargas, L. G. (2016). Predictive analytics model for healthcare planning and scheduling. *European Journal of Operational Research, 253(1)*, 121-131. doi:10.1016/j.ejor.2016.02.017

Hashem, I. A., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems, 47*, 98-115. doi:10.1016/j.is.2014.07.006

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, A. L. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggetions for research and applications. *International Journal of Production Economics, 154*, 72-80. doi:10.1016/j.ijpe.2014.04.018

Hazen, B. T., Skipper, J. B., Boone, C. A., & Hill, R. R. (2016). Back in business: operations research in support of big data analytics for operations and supply chain management. *Annals of Operations Research*, 1-11. doi:10.1007/s10479-016-2226-0

Hechtkopf, B. (2013, May 31). *How financial institutions can use big data and converged data sets to improve customer loyalty.* Retrieved April 8, 2017, from http://www.banktech.com/data-and-analytics/how-financial-institutions-can-use-big-data-and-converged-data-sets-to-improve-customer-loyalty/a/d-id/1296390?

Heinrich, A., Lojo, A., Rodríguez González, A., Vasiljevs, A., Garattini, C., Costa-Soria, C., . . . Abedjan, Z. (2016). *Big Data Technologies in Healthcare - Needs, opportunities and challenges.* Big Data Value Association.

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., . . . Zwitter, A. (2017, February 25). *Will Democracy Survive Big Data and Artificial Intelligence?* Retrieved May 9, 2017, from Scientific American: https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/

Helles, R., & Jensen, K. B. (2013). Introduction to the special issue: Making Data - Big Data and beyond. *First Monday, 18(10)*. doi:10.5210/fm.v18i10.4860

Henderson, J. (2012, May 29). *The Marketing Transformation at First Tennessee Bank.* Retrieved April 8, 2017, from IBM: https://www.ibm.com/blogs/watson-customer-engagement/2012/05/29/the-marketing-transformation-at-first-tennessee-bank/

Henry, C. (2016, January 14). *How Barclays is cashing in on big data & Hadoop to stay ahead in fintech.* Retrieved April 5, 2017, from http://www.cbronline.com/news/big-data/how-barclays-is-cashing-in-on-big-data-hadoop-to-stay-ahead-in-fintech-4785259/

Hillier, F. S., & Lieberman, G. J. (2015). *Introduction to Operations Research.* New York: McGraw-Hill Education.

Hitzler, P., & Janowicz, K. (2013). Linked Data, Big Data, and the 4th Paradigm. *Semantic Web 4*, 233-235.

Hmida, H. B., & Braun, A. (2016). Enabling an Internet of Things Framework for Ambient Assisted Living. In R. Wichert, & B. Mand, *Ambient Assisted Living - Advanced Technologies and Societal Change* (pp. 181-196). Cham, Zwitzerland: Springer International Publishing. doi:10.1007/978-3-319-52322-4_13

Honorato, M. (2016, March 31). *How your company can use big data analysis to improve last mile deliveries.* Retrieved April 14, 2017, from Beetrack: https://www.beetrack.com/en/blog/big-data-analysis-in-last-mile-deliveries

Howard, P. (2015, June 13). *Data Share: How big is the IoT?* Retrieved May 1, 2017, from http://philhoward.org/data-share-how-big-is-the-iot/

Hsieh, H.-C., Chang, K.-D., Wang, L.-F., Chen, J.-L., & Chao, H.-C. (2016). ScriptIoT: A Script Framework for and Internet-of-Things Applications. *IEEE Internet of Things Journal, 3(4)*, 628-636. doi:10.1109/jiot.2015.2483023

Hu, X., & Liu, H. (2012). Text analytics in social media. In C. C. Aggarwal, & C. Zhai, *Mining Text Data* (pp. 385-414). Springer US. doi:10.1007/978-1-4614-3223-4_12

IBM. (2015, December 15). *Durham Police Department - Reduces crimes with the help of IBM Big Data technology.* Retrieved April 23, 2017, from http://www-03.ibm.com/software/businesscasestudies/hk/en?synkey=E906175Y75689V95

IDC. (2016, May 23). *Worldwide Big Data and Business Analytics Revenues Forecast to Reach $187 Billion in 2019, According to IDC.* Retrieved May 1, 2017, from https://www.idc.com/getdoc.jsp?containerId=prUS41306516

IPCC. (2014). Climate Change 2014: Impacts, adaptation and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climat change. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, E. T. Bilir, . . . L. L. White (Eds.). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Isbitts, M. (n.d.). *Preventing Health Care Fraud with Big Data and Analytics.* Retrieved April 12, 2017, from LexisNexis Risk Solutions: http://www.lexisnexis.com/risk/insights/health-care-fraud-layered-approach.aspx

Ittmann, H. W. (2015). The impact of big data and business analytics on supply chain management. *Journal of Transport and Supply Chain Management, 9(1)*, 1-8. doi:10.4102/jtscm.v9i1.165

Jacobs, E. (2017, February 21). *8 veelvoorkomende misverstanden over de GDPR.* Retrieved April 27, 2017, from LegalWorld: http://www.legalworld.be/legalworld/8-veelvoorkomende-misverstanden-over-de-GDPR-2017.html?LangType=2067

Jha, S. (2016, October 31). *Apollo Hospitals uses big data analytics to control Hospita Acquired Infections.* Retrieved April 13, 2017, from ETCIO.com: http://cio.economictimes.indiatimes.com/news/case-studies/apollo-hospitals-uses-big-data-analytics-to-control-hospital-acquired-infections/55023676

Jin, J., Gubbi, J., Marusic, S., & Palaniswami, M. (2014). An Information Framework of Creating a Smart City through Internet of Things. *IEEE Internet of Things Journal, 1(2)*, 112-121. doi:10.1109/jiot.2013.2296516

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations & Production Management, 37(1)*, 10-36. doi:10.1108/IJOPM-02-2015-0078

Kahreh, M. S., Tive, M., Babania, A., & Hesan, M. (2014). Analyzing the applications of customer lifetime value (CLV) based on benefit segmentation for the banking sector. *Procedia - Social and Behavioral Sciences, 109*, 590-594. doi:10.1016/j.sbspro.2013.12.511

Kamienski, C., Jentsch, M., Eisenhauer, M., Kiljander, J., Ferrera, E., Rosengren, P., . . . Sadok, D. (2017). Application development for the Internet of Things: A context-aware mixed criticality systems development platform. *Computer Communications, 104*, 1-16. doi:10.1016/j.comcom.2016.09.014

Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly, 30(4)*, 508-513. doi:10.1016/j.giq.2013.05.012

Katal, A., Wazid, M., & Goudar, R. H. (2013). Big Data: Issues, Challenges, Tools and Good Practices. *Sixth International Conference on Contemporary Computing* (pp. 404-409). Noida, India: IEEE. doi:10.1109/IC3.2013.6612229

Katz, J. (2016, June 1). *Waarom big data in de zorg niet van de grond komt.* Retrieved April 12, 2017, from MKB Nederland: https://www.mkb.nl/forum/waarom-big-data-de-zorg-niet-van-de-grond-komt

Kayyali, B., Knott, D., & Van Kuiken , S. (2013, April). *The big-data revolution in US health care: Accelerating value and innovation.* Retrieved April 13, 2017, from McKinsey & Company - Healthcare Systems & Services: http://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-big-data-revolution-in-us-health-care

Kearn, M. (2016, March 1). *Machine Learning is for Muggles too!* Retrieved April 22, 2017, from https://blogs.msdn.microsoft.com/martinkearn/2016/03/01/machine-learning-is-for-muggles-too/

Khoury, M. J., & Ioannidis, J. P. (2014). Big data meets public health - Human well-being could benefit from large-scale data if large-scale noise is minimized. *Science, 346(1054)*, 1054-1055. doi:10.1126/science.aaa2709

Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-Data Applications in the Government Sector. *Communications of the ACM, 57(3)*, 78-85. doi:10.1145/2500873

Kim, S.-Y., Jung, T.-S., Suh, E.-H., & Hwang, H.-S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications, 31*, 101-107.

Kirchner, B. (2012, May 19). *Crunching the numbers.* Retrieved April 8, 2017, from The Economist: http://www.economist.com/node/21554743

Kiser, M. (2016, August 11). *Introduction to Natural Language Processing (NLP) 2016.* Retrieved May 11, 2017, from Algorithmia: http://blog.algorithmia.com/introduction-natural-language-processing-nlp/

Kobbacy, K. A., Vadera, S., & Rasmy, M. H. (2007). AI and OR in management of operations: history and trends. *Journal of the Operational Research Society, 58(1)*, 10-28. doi:10.1057/palgrave.jors.2602132

Konrad, R. A., Trapp, A. C., Palmbach, T. M., & Blom, J. S. (2017). Overcoming human trafficking via operations research and analytics: Opportunities for methods, models, and applications. *European Journal of Operational Research, 259(2)*, 733-745. doi:10.1016/j.ejor.2016.10.049

Konstantinova, N. (2014, March 30). *Why is Big Data Becoming So Popular.* Retrieved October 6, 2016, from NLP People: https://nlppeople.com/why-is-bigdata-becoming-so-popular/

Krumholz, H. M. (2014). Big Data and New Knowledge in Medicine: The Thinking, Training, and Tools Needed for a Learning Health System. *Health Affairs, 33(7)*, 1163-1170. doi:10.1377/hlthaff.2014.0053

Kumar, A. (2015). *Whitepaper - Omnichannel banking: A win-win proposition.* Retrieved April 9, 2017, from Infosys: https://www.infosys.com/industries/financial-services/white-papers/Documents/omni-channel-banking.pdf

Kumar, H., Singh, M. K., & Gupta, M. P. (2016). Smart governance for smart cities: a conceptual framework from social media practices. *Lecture Notes in Computer Science*, 628-634. doi:10.1007/978-3-319-45234-0_56

Laalaoui, Y., & Bouguila, N. (2015). *Artificial Intelligence Applications in Information and Communication Technologies.* Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-19833-0

Lafuente, G. (2015). The big data security challenge. *Network Security*, 12-14. doi:10.1016/S1353-4858(15)70009-7

Lamberti, F., Manuri, F., Sanna, A., Paravati, G., Pezzolla, P., & Montuschi, P. (2014). Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance. *IEEE Transactions on Emerging Topics in Computing, 2(4)*, 411-421. doi:10.1109/tetc.2014.2368833

Laney, D., & Jain, A. (2016, March 24). *100 Data and Analytics Predictions through 2020.* Retrieved May 1, 2017, from Gartner: https://www.gartner.com/doc/3263218?ref=clientFriendlyURL

Lantz, B. (2013a). *Machine Learning with R.* Birmingham, Uk: Packt Publishing Ltd.

Lantz, B. (2013b). Steps to apply machine learning to your data. In B. Lantz, *Machine Learning with R - Learn how to use R to apply powerful machine learning methods and gain an insight into real-world applications* (pp. 17-18). Birmingham: Packt Publishing.

Laurentian Bank. (2016, January 26). *Laurentian Bank Investor Forum - Our Transformation Plan: Becoming simpler, more efficient and more profitable [PowerPoint Slides].* Retrieved April 25, 2017, from https://www.laurentianbank.ca/pdf/20160124_LBC_Investor_Forum_008_web.pdf

Lawson, G., Salanitri, D., & Waterfield, B. (2016). Future directions for the development of virtual reality within an automotive manufacturer. *Journal of Applied Ergonomics, 53*, 323-330. doi:10.1016/j.apergo.2015.06.024

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in Big Data Analysis. *Science, 343(6176)*, 1203-1205. doi:10.1126/science.1248506

Lee, R. (2015). *Software Engineering, Artificial Intelligence, Networking and Parallel/Dsitributed Computing.* Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-10389-1

Lehmacher, W. (2016, March 31). *What is the scope of data mining in supply chain management?* Retrieved April 16, 2017, from Quora: https://www.quora.com/What-is-the-scope-of-data-mining-in-supply-chain-management

Leigh, D. (2010, November 28). *How 250,000 Us embassy cables were leaked.* Retrieved April 28, 2017, from The Guardian: https://www.theguardian.com/world/2010/nov/28/how-us-embassy-cables-leaked

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing, 80(6)*, 69-96. doi:10.1509/jm.15.0420

Levasseur, R. E. (2010). People Skills: Ensuring Project Success - A Change Management Perspective. *Interfaces, 40(2)*, 159-162. doi:10.1287/inte.1090.0473

Levasseur, R. E. (2015). People Skills: Building Analytics Decision Models That Managers Use - A Change Management Perspective. *Interfaces, 45 (4)*, 363-364. doi:10.1287/inte.2015.0798

Leveling, J., Edelbrock, M., & Otto, B. (2014). Big Data Analytics for Supply Chain Management. *International Conference on Industrial Engineering and Engineering Management* (pp. 918-922). IEEE. doi:10.1109/IEEM.2014.7058772

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., & Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C, 45*, 17-26. doi:10.1016/j.trc.2014.04.013

Liang, S. N., Tan, K. O., Clement, T. L., Ng, S. K., Mohammed, A. A., Musa, M., . . . Zulkifkli, Y. (2016). Open source hardware and software platform for robotics and artificial intelligence applications. *IOP Conference Series: Materials Science and Engineering, 114*, 1-10. doi:10.1088/1757-899X/114/1/012142

Liang, T.-P., Shen, K. N., & Guo, X. (2016). *Special Issue: Big Data Analytics for Business Intelligence*. Retrieved April 27, 2017, from Elsevier: https://www.journals.elsevier.com/expert-systems-with-applications/call-for-papers/special-issue-big-data-analytics-for-business-intelligence

Liang, Y. H. (2014). Customer Relationship Management and Big Data Mining. In W. Pedrycz, & S.-M. Chen, *Information Ganularity, Big Data, and Computational Intelligence* (pp. 349-360). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-08254-7_16

Liberatore, M. J., & Luo, W. (2010). The Analytics Movement: Implications for Operations Research. *Interfaces, 40(4)*, 313-324. doi:10.1287/inte.1100.0502

Liberatore, M., & Luo, W. (2013). ASP, The Art and Science of Practice: A Comparison of Technical and Soft Skill Requirements for Analytics and OR Professionals. *Interfaces, 43 (2)*, 194-197. doi:10.1287/inte.1120.0647

Liedtke, C. A. (2016). Research Report on "Quality, Analytics, and Big Data". *Strategic Improvement Systems, LLC*, 1-54.

Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications. *IEEE Internet of Things Journal*, 1-17. doi:10.1109/JIOT.2017.2683200

Lindsay. (2015, October 6). *Virtual Reality vs. Augmented Reality.* Retrieved May 2, 2017, from Augment: http://www.augment.com/blog/virtual-reality-vs-augmented-reality/

List, B. (2012, February 14). *Operations Research, Big Data to tie knot in April.* Retrieved March 20, 2017, from Informs: https://www.informs.org/About-INFORMS/News-Room/O.R.-Analytics-at-Work-Blog/Operations-Research-Big-Data-to-Tie-Knot-in-April

Liu, G., & Yang, H. (2017). Self-organizing network for variable clustering. *Annals of Operations Research*. doi:10.1007/s10479-017-2442-2

Liu, P., & Yi, S.-p. (2017). A study on supply chain investment decision-making and coordination in the Big Data environment. *Annals of Operations Research*. doi:10.1007/s10479-017-2424-4

Lizcano, P. E., Manchado, C., Gomez-Jauregui, V., & Otero, C. (2017). Virtual reality to assess visual impact in wind energy projects. In B. Eynard, V. Nigrelli, S. M. Oliveri, G. Peris-Fajarnes, & S. Rizzuti, *Advances on Mechanics, Design Engineering and Manufacturing* (pp. 717-725). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-45781-9_72

Lohrmann, D. (2014, January 12). *A mission for big data in government: How hyper-personalization can transform customer service.* Retrieved April 22, 2017, from Government Technology: http://www.govtech.com/blogs/lohrmann-on-cybersecurity/-A-mission-for-big-data-in-government-How-hyperpersonalization-can-transform-customer-service.html

Lombardi, P. L., Giordano, S., Farouh, H., & Yousef, W. (2012). Modelling the smart city performance. *Innovation the European Journal of Social Science Research, 25(2)*, 137-149. doi:10.1080/13511610.2012.660325

Longoria, G. (2015, August 4). *How the Internet of Things will shape the datacenter of the future.* Retrieved April 29, 2017, from Forbes Magazine: https://www.forbes.com/sites/moorinsights/2015/08/04/how-the-internet-of-things-will-shape-the-datacenter-of-the-future/#2eba6b0e2bf1

Lonzer, J. (2015, August 4). *Asthma management program uses big data to help Louisville residents breathe a bit easier.* Retrieved April 12, 2017, from Innovatemedtec: https://innovatemedtec.com/content/asthma-management-program-uses-big-data-to-help-louisville-residents-breathe-a-bit-easier

Lopez, D., Gunasekaran, M., Kaur, H., & Abbas, K. M. (2014). Spatial big data analytics of influenza epidemic in Vellore, India. *International Conference on Big Data* (pp. 19-24). IEEE. doi:10.1109/BigData.2014.7004422

Lord, N. (2017, January 27). *The History of Data Breaches.* Retrieved April 29, 2017, from Digital Guardian: https://digitalguardian.com/blog/history-data-breaches

Lv, Z., Song, H., Basanta-Val, P., Steed, A., & Jo, M. (2017). Next-Generation Big Data Analytics: State of the Art, Challenges, and Future Research Topics. *IEEE Transactions on Industrial Informatics*, 1-9. doi:10.1109/tii.2017.2650204

Lycett, M. (2013). 'Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems, 22(4)*, 381-386. doi:10.1057/ejis.2013.10

Maarala, A. I., Su, X., & Riekki, J. (2017). Semantic Reasoning for Context-aware Internet of Things Applications. *IEEE Internet of Things Journal, 4(2)*, 461-473. doi:10.1109/jiot.2016.2587060

Maenhout, B., & Vanhoucke, M. (2013). An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems. *Omega, 41*, 485-499. doi:10.1016/j.omega.2012.01.002

Malighetti, P., Paleari, S., & Redondi, R. (2009). Pricing strategies of low-cost airlines: the Ryanair case study. *Journal of Air Transport Management, 15*, 195-203. doi:10.1016/j.jairtraman.2008.09.017

Mancini, M. (2009, May). *Segmentation and Customer Loyalty - Using Segmentation to Strengthen Customer Loyalty.* Retrieved April 8, 2017, from nielsen: http://www.nielsen.com/content/dam/corporate/us/en/newswire/uploads/2009/06/segmentation-and-customer-loyalty-white-paper.pdf

Mansour, R. F. (2016). Understanding how big data leads to social networking vulnerability. *Computers in Human Behavior, 57*, 348-351. doi:10.1016/j.chb.2015.12.055

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011, May). *Big data: The next frontier for innovation, competition, and productivity.* Retrieved April 11, 2017, from McKinsey & Company: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation

Manzoni, J. (2017, February 21). *Big data in government: the challenges and opportunities.* Retrieved April 25, 2017, from gov.uk: https://www.gov.uk/government/speeches/big-data-in-government-the-challenges-and-opportunities

Marakas, G. M. (1999). *Decision support systems in the twenty-first century.* New Jersey, USA: Prentice-Hall, Inc. .

Marcello, S. (2016, August 16). *When and how you should secure PII (Personally Identifiable Information).* Retrieved April 29, 2017, from OPIN: http://www.opin.com/secure-personally-identifiable-information-pii/

Marchi, J. (2013, July 23). *Three more reasons to like visual analytics.* Retrieved April 22, 2017, from SAS: http://blogs.sas.com/content/sascom/2013/07/23/three-more-reasons-to-like-visual-analytics/

Marr, B. (2014, November 25). *Big Data: The Key Skills Businesses Need.* Retrieved March 2, 2017, from https://www.linkedin.com/pulse/20141125074301-64875646-big-data-the-key-skills-businesses-need

Marr, B. (2015, May 18). *Big data at Caesars Entertainment - A one billion dollar asset?* Retrieved April 8, 2017, from Forbes Magazine: https://www.forbes.com/sites/bernardmarr/2015/05/18/when-big-data-becomes-your-most-valuable-asset/#69210501eefd

Marr, B. (2016a, May 24). *Big Data: A Game Changer in Healthcare.* Retrieved May 9, 2017, from Forbes Magazine: https://www.forbes.com/sites/bernardmarr/2016/05/24/big-data-a-game-changer-in-healthcare/#36f73479525b

Marr, B. (2016b, November 17). *Big Data at Tesco: Real time analytics at the UK grocery retail giant.* Retrieved April 8, 2017, from Forbes Magazine: https://www.forbes.com/forbes/welcome/?toURL=https://www.forbes.com/sites/bernardmarr/2016/11/17/big-data-at-tesco-real-time-analytics-at-the-uk-grocery-retail-giant/&refURL=https://www.google.be/&referrer=https://www.google.be/

Marr, B. (2016c, December 13). *Big Data in Healthcare: Paris hospitals predict admission rates using machine learning.* Retrieved April 13, 2017, from Forbes Magazine: https://www.forbes.com/sites/bernardmarr/2016/12/13/big-data-in-healthcare-paris-hospitals-predict-admission-rates-using-machine-learning/#1332955d79a2

Marr, B. (2016d, March 23). *What Everyone Should Know About Cognitive Computing.* Retrieved May 11, 2017, from Forbes Magazine: https://www.forbes.com/sites/bernardmarr/2016/03/23/what-everyone-should-know-about-cognitive-computing/#796e23405088

Matthias, O., Fouweather, I., Gregory, I., & Vernon, A. (2017). Making sense of Big Data - can it transform operations management? *International Journal of Operations & Production Management, 37(1)*, 37-55. doi:10.1108/IJOPM-02-2015-0084

Matthies, M., Giupponi, C., & Ostendorf, B. (2007). Environmental decision support systems: Current issues, methods and tools. *Environmental Modelling & Software, 22(2)*, 123-127. doi:10.1016/j.envsoft.2005.09.005

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A revolution that will transform how we live, work, and think.* London: John Murray.

Mazzocchi, F. (2015). Could Big Data be the end of theory in science? - A few remarks on the epistemology of data-driven science. *EMBO reports, 16(10)*, 1250-1255. doi:10.15252/embr.201541001

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Bus Rev, 90(10)*, 61-67.

McCauley, R., & Graves, B. (2017, January 30). *How Data Visualization is the Future of Information Sharing.* Retrieved May 8, 2017, from Gov Tech: http://www.govtech.com/fs/How-Data-Visualization-is-the-Future-of-Information-Sharing.html

McColl-Kennedy, J. R., Gustafsson, A., Jaakkola, E., Klaus, P., Radnor, Z., Perks, H., & Friman, M. (2015). Fresh perspectives on customer experience. *Journal of Services Marketing, 29(6-7)*, 430-435.

McDonald, D., & Kelly, U. (2012). *Value and benefits of text mining.* JISC.

McLay, L. A. (2013, January 16). *Big data and operations research.* Retrieved March 6, 2017, from https://punkrockor.com/2013/01/16/big-data-and-operations-research/

Méndez, F. (2015, September 17). *Practical examples of Big Data use.* Retrieved April 8, 2017, from BBVA: https://www.bbva.com/en/news/economy/computerstudies-sciences-and-development/practical-examples-of-big-data-use/

Meng, Z., Wu, Z., Muvianto, C., & Gray, J. (2017). A Data-Oriented M3M Messaging Mechanism for Industrial IoT Applications. *IEEE Internet of Things Journal, 4(1)*, 236-246. doi:10.1109/jiot.2016.2646375

Menon, V., & Sheth, P. (2016, April 7). *Big Data Analytics can be a game changer for healthcare fraud, waste, and abuse.* Retrieved April 12, 2017, from https://www.hfma.org/Content.aspx?id=47523

Metcalf, J., Keller, E. F., & Boyd, D. (2016, May 23). *Perspectives on Big Data, Ethics, and Society.* Retrieved April 28, 2017, from Council for Big Data, Ethics, and Society: http://bdes.datasociety.net/wp-content/uploads/2016/05/Perspectives-on-Big-Data.pdf

Meyer, C., McGuire, T., Masri, M., & Shaikh, A. W. (2013, October 22). *Four steps to turn big data into action.* Retrieved March 2, 2017, from Forbes Magazine: https://www.forbes.com/sites/mckinsey/2013/10/22/four-steps-to-turn-big-data-into-action/#46da69b84380

Michael, G. (2015). The Future of Artificial Intelligence: Benevolent or Malevolent? *Skeptic (Altadena), 20(1)*, 57-61.

Miriovsky, B. J., Shulman, L. N., & Abernethy, A. P. (2012). Importance of Health Information Technology, Electronic Health Records, and Continuously Aggregating Data to Comparative Effectiveness Research and Learning Health Care. *Journal of Clinical Ontology, 30*, 4243-4248. doi: 10.1200/JCO.2012.42.8011

Mishra, D., Gunasekaran, A., Papadopoulos, T., & Childe, S. J. (2016). Big Data and supply chain management: a review and bibliometric analysis. *Annals of Operations Research*. doi:10.1007/s10479-016-2236-y

Mitchell-Guthrie, P. (2015, December 15). *OR vs. data science vs. analytics: what's in a name?* Retrieved March 25, 2017, from LinkedIn Pulse: https://www.linkedin.com/pulse/vs-data-science-analytics-whats-name-polly-mitchell-guthrie

Monks, T. (2016). Operational research as implementation science: definitions, challenges and research priorities. *Implementation Science, 11(1)*, 1-10. doi:10.1186/s13012-016-0444-0

Monnappa, A. (2016, April 5). *Data Science vs. Big Data vs. Data Analytics.* Retrieved March 1, 2017, from Simplilearn: https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article

Moore, D. F. (2016). Basic principles of survival analysis. In D. F. Moore, *Applied Survival Analysis Using R* (pp. 11-24). Springer International Publishing. doi:10.1007/978-3-319-31245-3_2

Morabito, V. (2015). Big Data and Analytics for Government Innovation. In V. Morabito , *Big Data and Analytics* (pp. 23-45). Springer International Publishing Switzerland. doi:10.1007/978-3-319-10665-6_2

Mori, K., & Christodoulou, A. (2012). Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review, 32(1)*, 94-106. doi:10.1016/j.eiar.2011.06.001

Mortenson, M. J., Doherty, N. F., & Robinson, S. (2015). Operational research form Taylorism to Terabytes: A research agenda for the analytics age. *European Journal of Operational Research, 241(3)*, 583-595. doi:10.1016/j.ejor.2014.08.029

MSG team. (n.d.). *Designing and Developing Decision Support Systems.* Retrieved March 7, 2017, from Management Study Guide: https://www.managementstudyguide.com/designing-and-developing-decision-support-systems.htm

Müller, V. C. (2016). *Fundamental Issues of Artificial Intelligence.* Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-26485-1

Müller, V. C., & Bostrom, N. (2016). Chapter 33 - Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In V. C. Müller, *Fundamental Issues of Artificial Intelligence* (pp. 555-572). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-26485-1_33

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Journal of the American Medical Association, 309(13)*, 1351-1352. doi:10.1001/jama.2013.393

Nair, S. G., & Puri, P. (2015). Open Source Threat Intelligence System. *International Journal of Research, 2(4)*, 360-363.

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data, 2(1)*, 1-21. doi:10.1186/s40537-014-0007-7

Nash, K. S. (2012, March 7). *How Big Data can reduce Big Risk.* Retrieved April 15, 2017, from CIO from IDG: http://www.cio.com/article/2371591/business-intelligence/how-big-data-can-reduce-big-risk.html

Nedbank Ltd. (2015, October 28). *Nedbank's wins big at global innovation awards with SME tool.* Retrieved April 5, 2017, from https://www.nedbank.co.za/content/nedbank/desktop/gt/en/news/nedbankstories/nedbankupdates/2015/nedbank_s-wins-big-at-global-innovation-awards-with-sme-tool.html

Nederlandse Zorgautoriteit. (2014). *Rapport Onderzoek zorgfraude.* Retrieved from https://www.nza.nl/1048076/1048181/Rapport_Onderzoek_zorgfraude__update.pdf

Nelson, P. (2015, July 10). *New sound monitoring by IoT can predict mechanical failure.* Retrieved April 14, 2017, from Network World: http://www.networkworld.com/article/2946195/internet-of-things/new-sound-monitoring-by-iot-can-predict-mechanical-failure.html

Nelson, R. (2015, August 18). *Why Netflix is winning the big data game.* Retrieved April 8, 2017, from YoungUpStarts: http://www.youngupstarts.com/2015/08/18/why-netflix-is-winning-the-big-data-game/

Novotny, M., Lacko, J., & Samuelcik, M. (2013). Applications of Multi-touch Augmented Reality System in Education and Presentation of Virtual Heritage. *Procedia Computer Science, 25*, 231-235. doi:10.1016/j.procs.2013.11.028

Nunan, D., & Di Domenico, M. (2013). Market research & the ethics of big data. *International Journal of Market Research, 55(4)*, 505-520.

Oaks, J. (2015, October 13). *T-Mobile mines Big Data and continues to progress.* Retrieved April 5, 2017, from http://www.smartdatacollective.com/jessoaks11/351284/t-mobile-mines-big-data-continues-progress

O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. (2015). Big data in manufacturing: a systematic mapping study. *Journal of Big Data, 2(1)*, 1-22. doi:10.1186/s40537-015-0028-x

OECD. (2010). Health care systems: Getting more value for money. *OECD Economics Department Policy Notes, 2*, 1-12.

OECD. (2013). *New Sources of Growth: Knowledge-Based Capital - Key Analyses and Policy Conclusions.* OECD Publishing. Retrieved from http://www.oecd.org/sti/inno/knowledge-based-capital-synthesis.pdf

Olshannikova, E., Ometov, A., Koucheryavy, Y., & Olsson, T. (2015). Visualizing Big Data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data, 2(1)*, 1-27. doi:10.1186/s40537-015-0031-2

Olson, D. L. (2015). The Use of Big Data and Data Mining in Supply Chains [PowerPoint slides]. 1-20. Retrieved from cbafiles.unl.edu/public/cbainternal/facStaffUploads/INFORMS2015.pptx

Ong, S. K., & Nee, A. Y. (2013). *Virtual and Augmented Reality Applications in Manufacturing.* Berlin: Springer Science & Business Media. doi:10.1007/978-1-4471-3873-0

O'Reilly Media. (2015, January 1). *Big Data Now: 2014 Edition*. Retrieved January 31, 2017, from O'Reilly: https://www.oreilly.com/ideas/big-data-now-2014-edition

Ottenheijm, S. (2015). *Big data in de gezondheidszorg - Definitie, toepassingen en uitdagingen.* Den Haag: TendITion. Retrieved from https://www.nictiz.nl/SiteCollectionDocuments/Rapporten/Big_data_in_de_gezondheidszorg.pdf

Oztekin, A. (2017). Creating a marketing strategy in healthcare industry: a holistic data analytic approach. *Annals of Operations Research*. doi:10.1007/s10479-017-2493-4

Pahl, J., Voss, S., & Sebastian, H.-J. (2017). Intelligent decision support and big data for logistics and supply chain management - A biased view. *Proceedings of the 50th Hawaii International Conference on System Sciences*, (pp. 1338-1340). Kona, United States.

Pal, K. (2015a, August). *Big Data Influence on Data Driven Advertising.* Retrieved April 3, 2017, from KDnuggets: http://www.kdnuggets.com/2015/08/big-data-influencing-data-driven-advertising.html

Pal, K. (2015b, October 27). *Use big data to improve the patient experience.* Retrieved April 13, 2017, from DataInformed: http://data-informed.com/use-big-data-to-improve-the-patient-experience/

Paladin. (2016). *How to Use Big Data to Achieve Marketing Goals: Combine your data, your people and your vision to make the right decisions for your business.* Paladin. Retrieved April 18, 2017, from http://www.paladinstaff.com/white-papers/download/PAL-The-Big-Data-Opportunity-whitepaper.pdf

Panettieri, J. (2017, February 9). *Cloud Market Share 2016: Amazon AWS, Microsoft Azure, IBM, Google.* Retrieved March 25, 2017, from Channel E2E: https://www.channele2e.com/2017/02/09/cloud-market-share-2017-amazon-microsoft-ibm-google/

Passy, J. (2016, March 23). *From branches to big data: how fifth third is shifting resources.* Retrieved April 8, 2017, from American Banker: https://www.americanbanker.com/news/from-branches-to-big-data-how-fifth-third-is-shifting-resources

Patel, A. B., Birl, M., & Nai, U. (2012). Addressing Big Data Problem Using Hadoop and Map Reduce. *Nirma International Conference on Engineering (NUiCONE).* doi:10.1109/nuicone.2012.6493198

Pentaho. (2016). Retrieved from 2016 Big Data Trends: http://events.pentaho.com/2016-gartner-mq-bi-registration-ggl.html?utm_campaign=Branded_EMEA&ad_group=Branded_EMEA_-_Big_Data&leadsource=Google_Ads&keyword=%2Bpentaho%20%2Bbig%20%2Bdata&utm_medium=cpc&medium=Google_Search&utm_source=google&gclid=CjwKEAjwg

Pettey, C. (2016, June 9). *Welcome to the API Economy.* Retrieved April 18, 2017, from Gartner: http://www.gartner.com/smarterwithgartner/welcome-to-the-api-economy/

Philips. (2016, September 29). *Big data: de grote sprong voorwaarts in de gezondheidszorg.* Retrieved April 10, 2017, from Knack Trends: http://trends.knack.be/economie/trends-information-services/big-data-de-grote-sprong-voorwaarts-in-de-gezondheidszorg/article-infoservices-758453.html

Philipson, T. (2016, August 4). *Are you just a number? How new healthcare value frameworks miss the point.* Retrieved April 13, 2017, from Forbes Magazine: https://www.forbes.com/sites/tomasphilipson/2016/08/04/are-you-just-a-number-how-new-healthcare-value-frameworks-miss-the-point/2/#74114256488b

Poleto, T., de Carvalho, V. D., & Costa, A. C. (2014). The Roles of Big Data in the Decision-Support Process: An Empirical Investigation. In B. Delibašić, J. E. Hernández, J. Papathanasiou, F. Dargam, P. Zaraté, R. Ribeiro, . . . I. Linden, *Decision Support Systems V - Big Data Analytics for Decision Making* (Vol. 216, pp. 10-21). Springer, Cham.

Poppelaars, J. (2011, May 22). *Analytics and Operations Research: a practitioner's view.* Retrieved March 20, 2017, from OR at Work: http://john-poppelaars.blogspot.be/2011/05/analytics-and-operations-research.html

Poppelaars, J. (2013, September 13). *The impact of Operations Research on People, Business and Society.* Retrieved March 19, 2017, from OR at Work: http://john-poppelaars.blogspot.be/2013/09/the-impact-of-operations-research-on.html

Power, D. J. (2002). *Decision support systems: concepts and resources for managers.* Westport, Conn.: Quorum Books.

Power, D. J. (2014). Using 'Big Data' for analytics and decision support. *Journal of Decision Systems, 23(2)*, 222-228. doi:10.1080/12460125.2014.888848

Press, G. (2016, August 5). *IoT Mid-Year Update From IDC And Other Research Firms.* Retrieved January 31, 2017, from Forbes Magazine: http://www.forbes.com/sites/gilpress/2016/08/05/iot-mid-year-update-from-idc-and-other-research-firms/#29e2b312765f

Proffitt, B. (2012, April 12). *Big data analytics may detect infection before clinicians.* Retrieved April 12, 2017, from IT World: http://www.itworld.com/article/2729007/big-data/big-data-analytics-may-detect-infection-before-clinicians.html

PTC. (2015). *White Paper - Quantifying the Return On Investment (ROI).* Retrieved October 17, 2016, from https://www.ptc.com/~/media/Files/PDFs/IoT/Quantifying_Return_On_Investment.pdf

Rajgopal, J. (2004). Chapter 11.2: Principles and applications of Operations Research. In K. B. Zandin, *Maynard's Industrial Engineering Handbook, 5th edition* (pp. 11.27-11.44). McGraw-Hill.

Ramshaw, A. (2011, July 28). *Implementing trigger-based marketing to drive customer loyalty [PowerPoint Slides].* Retrieved April 9, 2017, from SlideShare: https://www.slideshare.net/aramshaw/implementing-triggerbasedmarketingtodrivecustomerloyalty

Ranyard, J., Fildes, R., & Hu, T.-I. (2015). Reassessing the scope of OR practice: The Influences of Problem Structuring Methods and the Analytics Movement. *European Journal of Operational Research 245(1)*, 1-13. doi:10.1016/j.ejor.2015.01.058

Rashidi, P., Cook, D. J., Holder, L. B., & Schmitter-Edgecombe, M. (2011). Discovering activities to recognize and track in a smart environment. *IEEE Transactions on Knowledge and Data Engineering, 23(4)*, 527-539. doi:10.1109/tkde.2010.148

Ray, P. P. (2017). A survey of IoT cloud platforms. *Future Computing and Informatics Journal*, 1-12. doi:10.1016/j.fcij.2017.02.001

ReachForce. (2016, October 21). *5 Big Data Marketing Trends That Are Slated to Dominate 2017.* Retrieved April 10, 2017, from http://www.reachforce.com/blog/5-big-data-marketing-trends-that-are-slated-to-dominate-2017/

Redman, T. C. (2013, December). *Data's Credibility Problem.* Retrieved April 12, 2017, from Harvard Business Review: https://hbr.org/2013/12/datas-credibility-problem

Reitano, V. (2011, September 30). *Capers Jones creates a software-risk assessor.* Retrieved April 15, 2017, from SD Times: http://sdtimes.com/capers-jones-creates-a-software-risk-assessor/

Renzi, C., Leali, F., Cavazzuti, M., & Andrisano, A. O. (2014). A review on artificial intelligence applications to the optimal design of dedicated and reconfigurable manufacturing systems. *International Journal of Advanced Manufacturing Technology, 72(1-4)*, 403-418. doi:10.1007/s00170-014-5674-1

Richey, R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of Big Data in the supply chain. *International Journal of Physical Distribution & Logistics Management, 46(8)*, 710-739. doi:10.1108/IJPDLM-05-2016-0134

Roberts, M. (2014, June 25). *5 ways big data will impact quality management.* Retrieved April 14, 2017, from Hertzler Systems Inc.: http://www.hertzler.com/2014/06/5-ways-big-data-will-impact-quality-management/

Rymaszewska, A., Helo, P., & Gunasekaran, A. (2017). IoT powered servitization of manufacturing - an exploratory case study. *International Journal of Production Economics*, 1-14. doi:10.1016/j.ijpe.2017.02.016

Sampaio, A. Z., & Martins, O. P. (2014). The application of virtual reality technology in the construction of bridge: The cantlieve and incremental launching methods. *Automation in Construction, 37*, 58-67. doi:10.1016/j.autcon.2013.10.015

Sawant, N., & Shah, H. (2013). *Big Data Application Architecture Q&A - A problem-solution approach.* Apress. doi:10.1007/978-1-4302-6293-0

Scarfi, M. (2012, June 28). *Social Media and the Big Data Explosion.* Retrieved April 25, 2017, from Forbes Magazine: https://www.forbes.com/sites/onmarketing/2012/06/28/social-media-and-the-big-data-explosion/#18815e756a61

Schmalstieg, D., & Höllerer, T. (2017). Augmented Reality: Principles and Practice. *Virtual Reality* (pp. 425-426). Los Angeles, CA, USA: IEEE.

Sen, S., Barnhart, C., Birge, J. R., Boyd, E. A., Fu, M. C., Hochbaum, D. S., . . . Zenios, S. A. (2014). *Operations Research - A Catalyst for Engineering Grand Challenges.* National Science Foundation.

Shee, Y.-P., Crompton, D., Richter, H., & Maehle, S.-P. (n.d.). *Whitepaper: Big data in banking for maketers - How to derive value from big data.* Retrieved April 5, 2017, from Evry: www.evry.com

Shi, K. (2014, May 5). *Big data privacy report: seizing opportunities, preserving values [PowerPoint Slides].* Retrieved April 23, 2017, from https://www.slideshare.net/KezhanSHI/big-data-privacyreportmay

Shibl, R., Lawley, M., & Debuse, J. (2013). Factors influencing decision support system acceptance. *Decision Support Systems, 54(2)*, 953-961. doi:10.1016/j.dss.2012.09.018

Shneiderman, B. (2014). The Big Picture for Big Data: Visualization. *Science, 343(6172)*, 729-730. doi:10.1126/science.343.6172.730-a

Shumaker, R., & Lackey, S. (2014). Virtual, Augmented and Mixed Reality - Applications of Virtual and Augmented Reality. *16th International Conferenc on Human-Compute Interaction* (pp. 426-434). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-07464-1

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks, 76*, 146-164. doi:10.1016/j.comnet.2014.11.008

Sigaud, O., & Buffet, O. (2013). *Markov Decision Processes in Artificial Intelligence.* Hoboken, New Jersey, USA: John Wiley & Sons, Inc. doi:10.1002/9781118557426

Simoudis, E. (2016, March 14). *What's next for big data applications?* Retrieved March 7, 2017, from O'Reilly: https://www.oreilly.com/ideas/whats-next-for-big-data-applications

Simpson, D., Meredith, J., Boyer, K., Dilts, D., Ellram, L. M., & Leong, G. K. (2015). Professional, research, and publishing trends in Operations and Supply Chain Management. *Journal of Supply Chain Management, 51(3)*, 87-100.

Singh, D., Tripathi, G., & Jara, A. J. (2014). A survey of Internet-of-Things: Future Vision, Architecture, Challenges and Services. *World forum on Internet of Things (WF-IoT) 2014* (pp. 287-292). Seoul, South-Korea: IEEE.

Singh, H. (2016). Big Data and Natural Language Processing came together for better information extraction: Text Analytics. *International Journal of Scientific research and management, 4(11)*, 4902-4906. doi:10.18535/ijsrm/v4i11.15

Sodhi, M. S., & Son, B.-G. (2008). ASP, The Art and Science of Practice: Skills Employers Want from Operations Research Graduates. *Interfaces, 38(2)*, 140-146. doi:10.1287/inte.1080.0342

Sodhi, M., & Son, B.-G. (2010). Content analysis of OR job advertisements to infer required skills. *Journal of the Operational Research Society, 61*, 1315-1327. doi:10.1057/jors.2009.80

Someh, I. A., Wixom, B., Davern, M., & Shanks, G. (2017). Enablers and Mechanisms: Practices for Achieving Synergy with Business Analytics. *Proceedings of the 50th Hawaii International Conference on System Sciences*, (pp. 5358-5367).

Song, M.-L., Fisher, R., Wang, J.-L., & Cui, L.-B. (2016). Environmental performance evaluation with big data: theories and methods. *Annals of Operations Research*. doi:10.1007/s10479-016-2158-8

Splunk. (2016). *UniCredit delivers omni-channel excellence with real-time operational insights.* Retrieved April 9, 2017, from https://www.splunk.com/pdfs/customer-success-stories/splunk-at-unicredit.pdf

Stackowiak, R., Mantha, V., & Licht, A. (2016). *Improving insurer performance with Big Data - Architect's guide and reference architecture introduction.* Oracle Enterprise.

Stankovic, J. A. (2014). Research Directions for the Internet of Things. *IEEE Internet of Things Journal, 1(1)*, 3-9. doi:10.1109/JIOT.2014.2312291

Stokjoska, B. L., & Trivodaliev, K. V. (2017). A review of Internet of Things for smart home: Challenges and solutions. *Journal of Cleaner Production, 140(3)*, 1454-1464. doi:10.1016/j.jclepro.2016.10.006

Story, V., O'Malley, L., & Hart, S. (2011). Roles, role performance, and radical innovation competences. *Industrial Marketing Management, 40(6)*, 952-966. doi:10.1016/j.indmarman.2011.06.025

Strohm, C. (2016, October 13). *Predicting Terrorism from Big Data challenges U.S. Intelligence.* Retrieved April 23, 2017, from Bloomberg Technology: https://www.bloomberg.com/news/articles/2016-10-13/predicting-terrorism-from-big-data-challenges-u-s-intelligence

Sumathi, S., & Sivanandam, S. N. (2006). Introduction to Data Mining Principles. *Studies in Computational Intelligence, 29*, 1-20. doi:10.1007/978-3-540-34351-6_1

Sun, J., & Reddy, C. K. (2013). Big Data Analytics for Healthcare. *SIAM International Conference on Data Mining*, (pp. 1-112). Austin, TX, USA. Retrieved from https://www.siam.org/meetings/sdm13/sun.pdf

Sun, J., Wang, F., Hu, J., & Edabollahi, S. (2012). Supervised patient similarity measure of heterogeneous patient records. *SIGKDD Explorations, 14(1)*, 16-24. doi:10.1145/2408736.2408740

Suthakar, U., Magnoni, L., Smith, D. R., Khan, A., & Andreeva, J. (2016). An efficient strategy for the collection and storage of large volumes of data for computation. *Journal of Big Data, 3(1)*, 1-17. doi:10.1186/s40537-016-0056-1

Tan, J., & Kim, S. (2016). Applied Data Visualization in Virtual Reality. 1-7. Retrieved from http://web.stanford.edu/class/cs448b/cgi-bin/wiki-sp16/images/7/7a/Applied-data-visualization.pdf

Tena, M. (2016, June 28). *Data that help make better decisions.* Retrieved April 5, 2017, from BBVA: https://www.bbva.com/en/news/science-technology/technologies/data-that-help-make-better-decisions/

Tene, O. (2013, August 20). *Privacy and Big Data: The biggest public policy challenge of our time?* Retrieved April 28, 2017, from Privacy Perspectives: https://iapp.org/news/a/privacy-and-big-data-making-ends-meet/

Teradata. (2017). *The Future of Big Data.* Retrieved April 27, 2017, from http://bigdata.teradata.com/US/Big-Ideas/The-Future-of-Big-Data/

Tham, I. (2017, March 27). *Big data spells big opportunities here.* Retrieved May 1, 2017, from The Straits Times: http://www.straitstimes.com/tech/big-data-spells-big-opportunities-here

Toma, A. (2010). *Data Mining and Exploration.* Laval University.

Torra, V., & Narukawa, Y. (2015). Modeling Decisions for Artificial Intelligence. *12th International Conference on MDAI.* Skövde, Sweden: Springer International Publishing.

Tran, M., & Pham, V. (2016, May 10). *Manufacturing downtime cost reduction with predictive maintenance.* Retrieved April 14, 2017, from Arimo: https://arimo.com/machine-learning/2016/manufacturing-downtime-cost-reduction-predictive-maintenance/

Turcu, C. (2013). Re-thinking sustainability indicators: Local perspectives of urban sustainability. *Journal of Environmental Planning and Management, 56(5)*, 695-719. doi:10.1080/09640568.2012.698984

Tyler, N. (2016, December 13). *Tracksure sensor system to transform rail networks.* Retrieved April 14, 2017, from NewElectronics: http://www.newelectronics.co.uk/electronics-technology/tracksure-sensor-system-to-transform-rail-networks/149272/

United Nations. (2014). *World Urbanization Prospects - The 2014 Revision.* New York: United Nations.

Van Horenbeek, A., & Pintelon, L. (2013). A dynamic predictive maintenance policy for complex multi-component systems. *Reliability Engineering and System Safety, 120*, 39-50. doi:10.1016/j.ress.2013.02.029

Van Nieuwenhove, H. (2012, January 24). *Intego viert 15 jaar patiëntenregistraties.* Retrieved April 10, 2017, from Nieuws KULeuven: http://nieuws.kuleuven.be/node/7928

van Rijmenam, M. (2015, December 1). *7 Important Big Data Trends for 2016.* Retrieved April 25, 2017, from Datafloq: https://datafloq.com/read/7-big-data-trends-for-2016/1699

Vasudevan, S. (2016, November 22). *How HDFD Bank accelerated its digital push with CloudCherry's customer experience platform?* Retrieved April 9, 2017, from Dataquest: http://www.dqindia.com/how-hdfc-bank-accelerated-its-digital-push-with-cloudcherrys-customer-experience-management-platform/

Venkatesh, A. N. (2017). Connecting the Dots: Internet of Things and Human Resource Management. *American International Journal of Research in Humanities, Arts and Social Sciences*, 21-24.

Vera-Baquero, A., Colomo-Palacios, R., Molloy, O., & Elbattah, M. (2015). Business process improvement by means of Big Data based Decision Support Systems: a case study on call centers. *International Journal of Information Systems and Project Management, 3(1)*, 5-26. doi:10.12821/ijispm030101

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research, 261(2)*, 626-639. doi:10.1016/j.ejor.2017.02.023

Viechnicki, P., Khuperkar, A., Fishman, T. D., & Eggers, W. D. (2015, May 18). *Smart mobility; Reducing congestion and fostering faster, greener and cheaper transportation options.* Retrieved April 18, 2017, from Deloitte University Press: https://dupress.deloitte.com/dup-us-en/industry/public-sector/smart-mobility-trends.html

Vinod Kumar, T. M., & Dahiya, B. (2016). Chapter 1 - Smart Economy in Smart Cities. In T. M. Vinod Kumar, *Advances in 21st Century Human Settlements* (pp. 3-76). Singapore: Springer Nature. doi:10.1007/978-981-10-1610-3_1

Vlahogianni, E. I., Park, B. B., & van Lint, J. (2015). Big data in transportation and traffic engineering. *Transportation Research Part C: Emerging Technologies, 58*, 161. doi:10.1016/j.trc.2015.08.006

Wainewright, P. (2016, June 17). *How data hones conversations at scale at Moneysupermarket.* Retrieved April 9, 2017, from Diginomica: http://diginomica.com/2016/06/17/how-data-hones-conversations-at-scale-at-moneysupermarket/

Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics, 34(2)*, 77-84.

Walsh, C. (2015, November 9). *Understanding the data you have: a holistic approach to the data life cycle.* Retrieved April 13, 2017, from Excella Consulting: https://www.excella.com/insights/understanding-the-data-you-have-a-holistic-approach-to-the-data-life-cycle

Wang, C.-Y., Zhao, W., Liu, Q., & Chen, H.-W. (2017). Optimization of the tool selection based on big data. *Journal of Discrete Mathematical Sciences and Cryptography, 20(1)*, 341-360. doi:10.1080/09720529.2016.1183310

Wang, X., Kim, M. J., Love, P. E., & Kang, S.-C. (2013). Augmented Reality in built environment: Classification and implications for future research. *Automation in Construction, 32*, 1-13. doi:10.1016/j.autcon.2012.11.021

Wang, Y., Kung, L., & Byrd, T. A. (2016). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting & Social Change*, 1-11. doi:10.1016/j.techfore.2015.12.019

Ward, M. J., Marsolo, K. A., & Froehle, C. M. (2014). Applications of business analytics in healthcare. *Business Horizons, 57(5)*, 571-582. doi:10.1016/j.bushor.2014.06.003

Warth, J., Kaiser, G., & Kügler, M. (2011). The impact of data quality and analytical capabilities on planning performance: insights from the automotive industry. *10th International Conference on Wirtschaftsinformatik* (pp. 322-331). Zurich, Switzerland: Association for Information Systems Electronic Library.

Wauters, M., & Vanhoucke, M. (2017). A Nearest Neighbour extension to project duration forecasting with Artificial Intelligence. *European Journal of Operational Research, 259*, 1097-1111. doi:10.1016/j.ejor.2016.11.018

Wedel, M., & Kannan, P. K. (2016). Marketing Analytics for Data-Rich Environments. *Journal of Marketing, 80*, 97-121. doi:10.1509/jm.15.0413

Weinstein, L. S. (2015, November). *Innovations in London's transport: Big Data for a better customer experience [PowerPoint slides].* Retrieved from http://2015.data-forum.eu/sites/default/files/1600-1640%20Weinstein_SEC.pdf

Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *Journal of the American Medical Informatics, 20*, 144-151. doi:10.1136/amiajnl-2011-000681

Whitmore, A., Agarwal, A., & Xu, L. D. (2015). The Internet of Things - A survey of topics and trends. *Information Systems Frontiers, 17(2)*, 261-274. doi:10.1007/s10796-014-9489-2

Whitten, S. (2016, April 6). *Starbucks knows how you like your coffee.* Retrieved April 5, 2017, from http://www.cnbc.com/2016/04/06/big-data-starbucks-knows-how-you-like-your-coffee.html

Williamson, B. (2015). Educating the smart city: Schooling smart citizens through computational urbanism. *Big Data & Society, 2(2)*, 1-13. doi:10.1177/2053951715617783

Winig, L. (2016, February 18). GE's big bet on data and analytics. *MIT Sloan Management Review, 57*. Retrieved April 9, 2017, from MIT Sloan Management Review: https://sloanreview.mit.edu/case-study/ge-big-bet-on-data-and-analytics/

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). *Data Mining - Practical Machine Learning Tools and Techniques (4th edition).* Cambridge, USA: Elsevier.

Woodie, A. (2016, April 15). *How Intuit personalizes TuboTax experiences with Big Data.* Retrieved April 5, 2017, from https://www.datanami.com/2016/04/15/intuit-personalizes-turbotax-experiences-big-data/

Wortmann, F., & Flüchter, K. (2015). Internet of Things - Technology and Value Added. *Business & Information Systems Engineering, 57(3)*, 221-224. doi:10.1007/s12599-015-0383-3

Wu, X., Zhu, X., Wu, G.-Q., & Ding, W. (2014). Data Mining with Big Data. *IEEE Transactions on knowledge and data engineering, 26(1)*, 97-107. doi:10.1109/tkde.2013.109

Xu, L. D., He, W., & Li, S. (2014). Internet of Things in Industries: A Survey. *IEEE Transactions on Industrial Informatics, 10(4)*, 2233-2243. doi:10.1109/tii.2014.2300753

Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research, 69(5)*, 1562-1566. doi:10.1016/j.jbusres.2015.10.017

Yang, X.-S. (2013). *Artificial Intelligence, Evolutionary Computing and Metaheuristics.* Berlin: Springer Berlin Heidelberg. doi:10.1007/978-3-642-29694-9

Yulinsky, C. (2012, January 24). Decisions, decisions... will 'Big Data' have 'Big' impact? *Financial Times*. Retrieved April 3, 2017, from https://www.ft.com/content/9ee048b6-4612-11e1-9592-00144feabdc0

Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of Things for Smart Cities. *IEEE Internet of Things Journal, 1(1)*, 22-32. doi:10.1109/JIOT.2014.2306328

Zang, Y., Zhang, F., Di, C.-a., & Zhu, D. (2015). Advanced of flexible pressure sensors toward artificial intelligence and health care applications. *Material Horizons, 2(2)*, 140-156. doi:10.1039/c4mh00147h

Zhan, Y., Tan, K. H., Li, Y., & Tse, Y. K. (2016). Unlocking the power of big data in new product development. *Annals of Operations Research,*. doi:10.1007/s10479-016-2379-x

Zhang, L. (2016). *Doctoral Thesis - Big Data Analytics for fault detection and its application in maintenance.* Lulea: Lulea University of Technology.

Zhang, Y., Shu, S., Ji, Z., & Wang, Y. (2015). A Study of the Commercial Application of Big Data of the International Hotel Group in China. *International conference on Big Data Computing Service and Applications* (pp. 412-417). IEEE.

Zhaparov, M. K., & Nassen, Y. (2016). 3D modelling based on virtual reality. *6th International Conference on Cloud System and Big Data Engineering* (pp. 399-402). Noida, India: IEEE.

Zibouh, O., Dalli, A., & Drissi, H. (2016). Cloud computing security through parallelizing Fully Homomorphic Encryption applied to multi-cloud approach. *Journal of Theoretical and Applied Information Technology, 87(2)*, 300-307.

Zwakman, G., Aslett, M., Stamper, J., Curtis, J., & Roy, K. (2016, june 14). *Total Data market expected to reach $132 bn by 2020.* Retrieved May 1, 2017, from 451 Research: https://451research.com/report-short?entityId=89339&referrer=marketing&utm_source=website_homepage&utm_medium=website&utm_term=data_platforms_analytics&utm_content=apply_for_trial&utm_campaign=2016_market_insight

Zweigenbaum, P., Demner-Fushman, D., Yu, H., & Cohen, K. B. (2007). Frontiers of biomedical text mining: current progress. *Briefings in bioinformatics, 8(5)*, 358-375. doi:10.1093/bib/bbm045