Politecnico di Torino

Dipartimento di Ingegneria Gestionale e della Produzione

Corso di Laurea Magistrale in Ingegneria Gestionale

Classe n. LM-31 – Ingegneria Gestionale

Big Data: Unleashing their economic value.
A literature review

Relatore:
Chiar.mo Prof. Paolo Neirotti

Correlatore:
Chiar.ma Prof.ssa Elisabetta Raguseo

Tesi di Laurea Magistrale di

Giampiero D'Arconso

Matricola n. S231699

Anno accademico 2017 / 2018

Index

1. Introduction

Big data is a term used to describe the massive amounts of information that companies, government organizations, and academic institutions can use to do anything (Krugler, 2016). It is also a large quantity of structured, semi structured or unstructured data continuously generated from diversified sources, which inundates business operations in real time and impacts on decision-making (Jie Sheng, 2017). It is emerged as an exciting frontier of productivity and opportunity in the last decade, giving the issue to understand "How organizations translate big data potential into economic value?". In a technological environment firms are trying to align their asset to reach even more competitive advantage on market. Digitalization most often refers to enhancing business models, business operations, business functions and activities, by leveraging digital assets with the aim to improve knowledge and gain a specific benefit. Among all innovative initiatives, big data has gained even more attention from industries and also academic researcher. International Data Corporation says that "worldwide revenues for big data and business analytics will grow from $130.1 billion in 2016 to more than $203 billion in 2020 and a much longer list of what are the effective amount of investments by firms are available in every journals. Given the importance and the impact by which this technology is appearing and transforming the organizations, this paper has the objective to explain how effectively companies could unleashing its potential value. As regard big data it is possible to describe the material resources, like what are the most common database architectures or program languages; what are the most dynamic and scalable algorithms, or what are the skills required from people. But all of these will be only mentioned. The core of the research is to understand what the effective implication are in manage big data, what are the desirable benefits and what the effective solutions. There is no implementation of cluster database using scalable fault tolerance software but a framework where prerequisites and consequences are collected and analyze.

This work is in the form of the literature review, that is a text of an academic paper, which represents the current knowledge of the research topic, including substantive findings, as well as theoretical and methodological contributions. The Literature Review proposed has the aim to collect all the previous big data works to understand

what the status of art is and to support future studies clarifying all the effect generated by this technology. So, the study examines the following research question 'What is the effect of Big data capability on firm performance?'. The work addresses this topic by introducing big data capability and consulting the literature on resourced based view, dynamic capability and other management related theories. Data driven decision making enables a firm to create dynamic knowledge that subsequently lead to competitive advantage. What follow is developed through four main blocks. It will be first a definition of what is big data both for academics and business organization with the tentative to provide a more accurate description. Then is described the research methodology to explain what are been the steps conducted to collect and analyze various papers. Then is developed a Big Data Framework to investigate the relations between big data capabilities and firm performances, dividing in two main levels that represent the key findings: (i) the need to own various organizational-level resources that in combination build a big data analytics capability; (ii) the identification of four main firm performances improved with the adoption of three big data capabilities. In addition, the effort to extend the framework with the introduction of mediator and moderator variable.

The last part of the thesis explores (I) the main big data initiatives adopted from companies in five industries: healthcare, public administration, manufacturing, retail and bank, the area in which these new technologies are getting more attention. And (II) the application of the framework across industries trying to underline what are the real example of big data articulation in the specific domain. All the work is concluded with a section that presents expected results of research and suggests for future investigation.

2. Big data definition and concept

In recent years there are different hot topics in scientific and technological fields. The explosion of innovation rate raises the necessity to understand and adopt these new tools to better shape own business. Among Internet of Things, Blockchain, Artificial intelligent there is a phenomenon of massive data creation that is giving even more attention: "big data". Big data is identified as the large volume of data, both structured and unstructured, that inundates the daily business. The degree of popularity of this phenomenon is accompanied with several inconsistent meanings and lacks formal definition (Andrea De Mauro, 2016). Big data term was used in 1990s-2000s probably due to works of Peter Denning, John Mashey and Doug Laney that in this decade contributed writing articles on analytics technique, process of massive amount of information. Data deluge (Khan, 2017), mass data, fourth paradigm and ocean of data are just few of the several ways to represent the object. Big data and the way in which is produced, introduced important changes in information flows and became relevant for organizations (Constantiou I.D., 2014). The excitation that big data term brings with it is due to the impressive consequential activities that born from their utilization. The necessity to manage this data leads to the so-called big data revolution, that represents a trigger for the design and development of the new models, methods, and systems for collecting and analysing consumers' behaviour (Abbasi A., 2016). Three major trends have allowed the explosion of big data interest. First mobile application, GPS sensors, logfiles, and social web are used every time and it has become much cheaper to generate a wide variety of data. Second, it has become easier to process large amounts of data, due to advances open source software, inexpensive cloud and qualified data scientist. Finally, data have become public, so available for everyone (Abadi D., 2016). In the entire world the 90% of data was generated during the past two years, with billions of data created every day. The U.S. President's Council of Advisors on Science and Technology recently identified Big Data as a "research frontier" that can "accelerate progress across a broad range of priorities". (Gang-Hoon Kim, 2014). And the list of implications related to the use of information seem not end, some of other curious reasons are summarized in **Table 1**. The literature review has explored the big data meaning specify according to the famous adjectives that are

associated to this trend. For this reason, a collection of academic definitions and an analysis on 5-V meaning is included in the next chapters.

Table 1: "20 Mind-Boggling Facts Everyone Must Read", **(Marr, 2015)**

1. "The data volumes are exploding; more data has been created in the past two years than in the entire previous history of the human race".
2. "Data is growing faster than ever before and by the year 2020, about 1.7 megabytes of new information will be created every second for every human being on the planet".
3. "By then, our accumulated digital universe of data will grow from 4.4 zettabytes today to around 44 zettabytes, or 44 trillion gigabytes".
4. "Every second we create new data. For example, we perform 40,000 search queries every second (on Google alone), which makes it 3.5 searches per day and 1.2 trillion searches per year".
5. "In Aug 2015, over 1 billion people used Facebook FB -0.52% in a single day"
6. "Facebook users send on average 31.25 million messages and view 2.77 million videos every minute".
7. "We are seeing a massive growth in video and photo data, where every minute up to 300 hours of video are uploaded to YouTube alone".
8. "In 2015, staggering 1 trillion photos will be taken and billions of them will be shared online. By 2017, nearly 80% of photos will be taken on smart phones".
9. "This year, over 1.4 billion smart phones will be shipped - all packed with sensors capable of collecting all kinds of data, not to mention the data the users create themselves".
10. "By 2020, we will have over 6.1 billion smartphone users globally (overtaking basic fixed phone subscriptions)".
11. "Within five years there will be over 50 billion smart connected devices in the world, all developed to collect, analyse and share data".
12. "By 2020, at least a third data will pass through the cloud (a network of servers connected over the Internet)".
13. "Distributed computing (performing computing tasks using a network of computers in the cloud) is very real. Google GOOGL -0.45% uses it every day to involve about 1,000 computers in answering a single search query, which takes no more than 0.2 seconds to complete".
14. "The Hadoop (open source software for distributed computing) market is forecast to grow at a compound annual growth rate 58% surpassing $1 billion by 2020".
15. "Estimates suggest that by better integrating big data, healthcare could save as much as $300 billion a year — that's equal to reducing costs by $1000 a year for every man, woman, and child".
16. "The White House has already invested more than $200 million in big data projects".
17. "For a typical Fortune 1000 company, just a 10% increase in data accessibility will result in more than $65 million additional net income".
18. Retailers who leverage the full power of big data could increase their operating margins by as much as 60%.
19. "73% of organizations have already invested or plan to invest in big data by 2016".
20. "And one of my favourite facts: At the moment less than 0.5% of all data is ever analysed and used, just imagine the potential here".

2.1. Big data in the literature

It is considered interesting paint a representations of big data term collecting academic, scientific and business work in order to analyse different connotation given for the different scope **Table 2**. From a preliminary analysis is a hard work to find a shift in meaning and use. Although even more technology tools, applications and actions are generated every day, the main concept still unaltered. Consulting company as Mckinsey, IBM give a detailed description of what big data effectively is (Manyika, 2011), while other academic studies are trying to resume the other abstract and deep aspects.

In this section is introduced a "competitor" of 5Vs definition: the 3Cs attributes. In particular after the analysis of different enouncements the three new or relevant big data concepts are: Complexity, Capital and Challenge. *Complexity* relates to the hard work of analysing, storing, filtering and utilizing the data (Addo-Tenkorang R., 2016). The huge amount of information handled by companies as banks, e-commerce is considered a *capit*al, a potential valuable asset, a resource that needs to be managed (Chasalow, 2015). *Challenge* started to be shaped recently with the agitation that firms have in realizing functional and with marginality programs (Abadi D., 2016).

Particularly interesting is the evolutions of BI&A during the years. The work of Chen H., 2012 individualized three macro steps in definition and composition of big data. Business analytics found its roots in 1950 and only in the last decade of XX century assume relevant connotation in companies. In Business Intelligence & Analytics 1.0 the data are mostly structured, collected by companies through various legacy systems, and often stored in commercial relational database management systems (RDBMS). Only with the diffusion of internet, web contents data assume hidden insights that require different algorithm for the analysis. This process, cantered on text and web analytics for unstructured data, gives a different connotation and represents the Business Intelligence &Analytics 2.0. But for the first time in 2011 new technologies surpassed PCs and laptops: mobile phones and tablets. This new tool creates information every second with the necessity of a real-time analysis. Data that came from this device were generated in different forms and required sophisticated tool to gain all their potentiality. Internet of things and artificial intelligence are representing

the new frontier of big data creating the Business Intelligence & Analytics 3.0 era (Chen, 2012).

Table 2: Big data definition across literature

References	Definition
(MANYIKA, 2011)	"Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze".
(BOYD D., 2012)	"Big Data as a cultural, technological, and scholarly phenomenon that rests on the interplay of: Technology, Analysis, Mythology".
(CHEN, 2012)	"The data sets and analytical techniques in applications that are so large (from terabytes to exabytes) and complex (from sensor to social media data) that they require advanced and unique data storage, management, analysis, and visualization technologies".
TECH AMERICA FOUNDATION'S FEDERAL BIG DATA COMMISSION (2012)	"Big data is a term that describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information."
(CONSTANTIOU I.D., 2014)	"Miscellaneous records of the whereabouts of large and shifting online crowds".
(JAGADISH H.V., 2014)	"Data is being collected at an unprecedented scale. Decisions that previously were based on guesswork, or on painstakingly handcrafted models of reality, can now be made using data-driven mathematical models".
(ABADI D., 2016)	"Big data as a defining challenge of our time. Big data arose due to the confluence of three major trends: much cheaper to generate, much cheaper to process and data management has become democratized".
(KRUGLER, 2016)	"Term used to describe the massive amounts of information companies, government organizations, and academic institutions can use to do, well, anything".
(ANDREA DE MAURO, 2016)	"Information assets characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value".
(ADDO-TENKORANG R., 2016)	"A fast-growing amount of data from various sources that increasingly poses a challenge to industrial organizations and also presents them with a complex range of valuable-use, storage and analysis issues".
(YAQOOB I., 2016)	"The set of structured, unstructured, and semi-structured data accumulated from heterogeneous data sources".
(CHEN D. Q., 2015)	"A new form of capital along with the traditional financial and intellectual capitals for businesses".
(KHAN, 2017)	"Data deluge, the excessively huge volume of data generated at a regularly increasing basis in the world".
(JIE SHENG, 2017)	"Extremely large amount of structured, semi structured or unstructured data continuously generated from diversified sources, which inundates business operations in real time and impacts on decision-making through mining insightful information from rambling data".
GARTNER IT GLOSSARY, (N.D.)	"Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making".

2.2. What are the V-aspects of Big data?

As introduced in **Table 2**, in the Gartner It Glossary: "Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" or similarly as Laney (2001) reports, Volume, Variety, and Velocity (or the Three V's) are the three dimensions of challenges in data management.

The scope of this chapter is to analyse in detail what does big data mean for science, healthcare, manufacturing and retail comparing with the V-adjective. In particular **Table 3** defines the aspects of volume, variety, velocity, veracity, value reporting also a specific example to enter more in detail in the issue. Precedent literature reviews catalogued references filtering V-actions in different aspects or sector. Addo-Tenkorang R., (2016) present an useful analysis of what big data represent for supply chain, but also more general presentation can be found due to positive implication on companies' business. (Wamba S. F. S. A., 2015).

Volume is the most common adjective attributed to big data. It is underlined also in the term itself. Volume gives to big data a quantitative connotation. Terabytes (10^{12}) and Petabytes (10^{15}) are the measure unit consulted to express the magnitude of data (Gandomi A., 2015). For example, retail market as Tesco and Walmart generate every month 1,5 billion on new items or store something like 2,5 petabytes of information in data warehouse (Manyika, 2011). But volume is also considered to be context dependent. The size of data could be not univocal and weighted. For this reason, it is not meaningful to define a specific range to measure the volume of big data in a universal sense. So, it is cast as "big" if the management, storage, retrieval and analysis with common technology and tools start to be difficult (Guoqing Chen, 2016).

Variety is inherent to the heterogeneities of the structures, formats, and sources of Data (Guoqing Chen, 2016). The source of information could be multiple. The main classification is unstructured and structured data. As regard the first definition, are considered all the information not evident that have the potentiality to be understood. A music record, video or a photo on web site represent an example of unstructured data transformed in knowledge with specific analysis and tools. The structured data are information ready for the analysis and that need only the storage and not the

conversion. User-generated contents can be used to assess and inform pricing decisions in near real time, for example Tata Motors analyses 4 million text messages every month, spanning everything from product complaints to reminders about service appointments to announcements about new models, as well as connecting these with customer satisfaction polling (Agarwal, 2012).

Velocity refers to the rate at which data are produced and yet the speed at which data should be retrieved and analysed by a big data analytics system. Application as smartphones, apps, are used every second generating a wave of information. For many applications, the speed of data creation is even more important than the volume. The real-time generation is translated in the necessity of high rate analysis that could be useful for the creation of personalized offers for everyday customers (Cukier K., 2010). This makes possible for a company to be much more agile than its competitors (Manyika, 2011). Giant like Wal-Mart, processes more than one million transactions per hour. On Facebook alone, we send 10 billion messages including photos and videos per day; we click the 'share' button 4.5 billion times and upload 350 million new pictures each and every day (Thibeault, 2014). The previous 3Vs are recently extended with other 2Vs: veracity and value. The fourth V, *veracity*, is concerned on quality and reliability, and assume connotation of objectivity, truthfulness and credibility of data. The generation process of information starts from different sources that a prior can't be trustable and controlled. Is defined as challenging also the technology to manage data (McAfee & Brynjolfsson, 2012). Think about the 20% of web contents that are classified as spam (Abbasi A. S. S., 2016) and generated from users and mobile applications utilized for casual actions. To ensure the quality of information especially in the public administration sector some organization are creating curious program: Politifact (albeit based on man-powered fact-checking) and TruthGoggles have the capability to filter true facts helping citizens to develop better understanding of politician statements (Rubin, 2012). As Davenport report eBay Inc. was in the middle of an enormous data replication problem but developing an internal website (datahub) which enables managers to filter data replication was able to realign the records. (Davenport, 2012). The last V added on Big Data definition leads to the question of *value*. In order to support decision with useful information the main task is to eliminate unimportant and irrelevant data, so that the remaining data are useful. There are

different economic values and good information hidden amongst a larger body of non-traditional data. The challenge of new software must be the rapid (timely analysis) identification of what is relevant (Dijcks, 2013). There are a vastity of valuable program launched by retail, e-commerce and manufacturing organization. In the car industry Ford captures primary consumer data from GPS sensors of four million of its vehicles on the road to develop innovation and create value added against competitor such as Hyundai (King, 2012).

Table 3: Definition and examples of 5Vs

Attributes	Nature	Examples
Volume	Volume refers to the magnitude of data. Big data sizes are reported in multiple terabytes and petabytes. (Gandomi A., 2015)	• Tesco generates more than 1.5 billion new items of data every month (Manyika, 2011) • Wal-Mart's data warehouse includes some 2.5 petabytes of information (Manyika, 2011)
Variety	Variety refers to the structural heterogeneity in a dataset. Technological advances allow firms to use various types of structured, semi-structured, and unstructured data. (Gandomi A., 2015)	• Tata Motors analyses 4 million text messages every month, spanning everything from product complaints to reminders about service appointments to announcements about new models, as well as connecting these with customer satisfaction polling (Agarwal S., 2014)
Velocity	Velocity refers to the rate at which data are generated and the speed at which it should be analyzed and acted upon. (Gandomi A., 2015)	• Wal-Mart processes more than one million transactions per hour. The data emanating from mobile devices and flowing through mobile apps produces torrents of information that can be used to generate real-time, personalized offers for everyday customers. (Cukier K., 2010) • On Facebook alone, we send 10 billion messages including photos and videos per day; we click the 'share' button 4.5 billion times and upload 350 million new pictures each and every day (Thibeault, 2014)
Veracity	Veracity represents the unreliability inherent in some sources of data. (Gandomi A., 2015)	• *Politifact* (albeit based on man-powered fact-checking) and *TruthGoggles* sort true facts in politics helping citizens to develop better understanding of politicians tatements (Rubin, 2012)
Value	The economic value of different data varies significantly. There is good information hidden amongst a larger body of non-traditional data (Dijcks, 2013);	• Ford captures primary consumer data from around four million of its vehicles on the road through sensors and remote app-management software to develop innovation and create value added against competitor such as Hyundai (King, 2012)

3. Review methodology

The chapter is divided in two paragraphs in which are explained the main principles and steps adopted for the research. Following the scheme of the literature review (Nguyen T., 2017), first a description of the conceptual model of researching is presented and then, in the second section, a report of the results is shown.

3.1. Research methodology

3.1.1. Literature review research method

The present work explores different areas in which Big Data has been discussed, from the application in various sectors such as manufacturing, retail etc. to the different technology tools. Given the novelty of Big Data Analytics in a management context, a great depth of knowledge is required for being able to give a contribute in line with the research question (Kache F., 2017).

The literature research is conducted considering journals listed by the Associations Business School (ABS) guide, in which it is possible filter subject field and relative quality of journals focused on business and management. There are approximately 22 "Subject Fields" and are classified according to their scientific quality from 1 to 4 (lower grade – higher grade). For this literature the focus is on those belonging to Information Management that better fit with the main scope of this work and journals with the highest ranking (3 and 4) **Figure 1**.

The literature review process is articulated in five steps, with two steps of pure research and three of filtering, elimination of duplication and analysis. Although the topic can be shaped in different ways, "business intelligence", "data analytics" and, "large dataset", the selecting criteria adopted in this work is search all the articles, literature reviews, conference papers and editorial with the term "Big Data" and "Data driven decision making" in the title, abstract and as keywords. This kind of limitation is common in bibliometric studies (Chen, 2012) but ensures data consistency and relevance across the collection. In addition, the same criteria are used to collect all the

literature reviews, consulting reports, and references present in Google scholar and ResearchGate, two of the major scientific platforms.

The term big data became widespread as recently as in 2011 due to IBM and other leading technology companies' investments in building the niche analytics market (Gandomi A., 2015). For this reason, the date range used specifically for this purpose is 2011-2017 to ensure a coverage of all the fields of application, to capture the probable shift of the term in utilization and meaning, and to understand the progress of the scientific studies on this theme. To refine the research a content analysis of the titles and abstracts is executed to extract paper with the same area that are not identified with the previous criteria (Halevi, 2012).

Name of Journal	ABS rating
MIS Quarterly	4
Information Systems Research	4
Communications of the ACM	3
IEEE Transactions on Software Engineering	3
Decision Support Systems	3
Journal of Information Technology	3
Information Processing and Management	3
Information and Management	3
Journal of Management Information Systems	3
Expert Systems with Applications	3
European Journal of Information Systems	3
Journal of Strategic Information Systems	3
International Journal of Human-Computer Studies	3
ACM Transactions on Computer-Human Interaction	3
International Journal of Electronic Commerce	3
INFORMS Journal on Computing	3
Information Systems Journal	3
Information and Organization	3
Journal of the Association of Information Systems	3
Journal of the American Society for Information Science and Technology	3

Figure 1: ABS journals and ranking

This comprehensive review of the articles whose topic deals with big data concept is conducted to entails three characteristics: (i) the development of a performances framework, (ii) an overview about big data technologies and, (iii) an understanding of big data's impact on various sectors. (Wamba S. F. S. A., 2015).

3.1.2. Research method for framework

The main aim of this study is to realize a Big data capabilities framework structured in three levels. The first level is referred to the big data components, classified through a Resourced Based view (Peteraf, 2003). In particular are individualized seven tangible, intangible and human resources: data (internal, external), basic resource (time, investments), technology (Hadoop, NoSql.), technical skills, managerial skills, organizational learning (ability to explore, store, share, and apply knowledge) and data driven culture. These are antecedents of three big data capabilities components: big data technological capability, big data management capability and big data human capability. The final ramification are the firm's performances: strategic performance, financial performance, marketing performance and operational performance.

The research conducted to realize the framework differs from the literature research, in the keywords choice. In this analysis are utilized: "big data framework", "quantitative model", "research model", "big data components" and "big data firm performance" as search sentences. The choice of more addressed words is due to the huge amount of works in the academic field. The same sources and databases are consulted. The results of the research are included in the quantitative analysis of the next section.

3.2. Results and selection

In recent years the term Big Data is giving attention. Every day thousands of web contents born, and different studies start with the aim to deeply understand its consequences and its benefits.

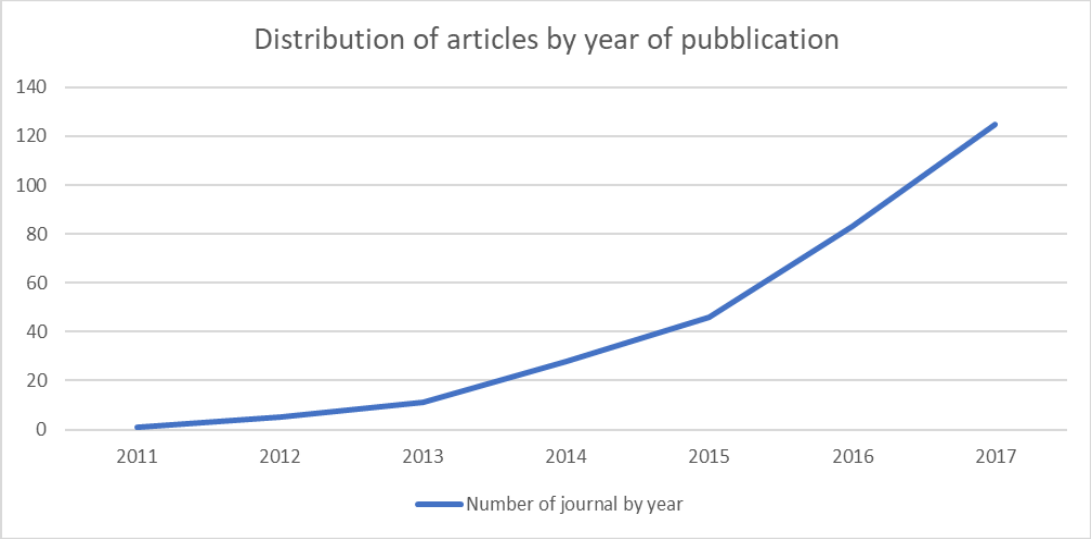

Figure 2: Distribution of articles by year

Tapping on Google "Big Data", 343M of links are collected and shown. Tv, journals have dedicated sections and every year conferences are organized. The result of the study points out the growth rate of interest in this topic (**Figure 2**) with an increase of publications during the year.

The research is conducted within the following databases: Elsevier (SCOPUS), Emerald, IEEE Xplore, ScienceDirect, and Taylor & Francis. The **Figure 3** collects the list of the number of articles in ABS journals. The amount is different according to the coverage of topics that the editor has planned. For example, Expert System with Application is a journal whose focus is on discovery intelligent systems applied in industry, government and other sectors, resulting one of the major source of works related to this theme.

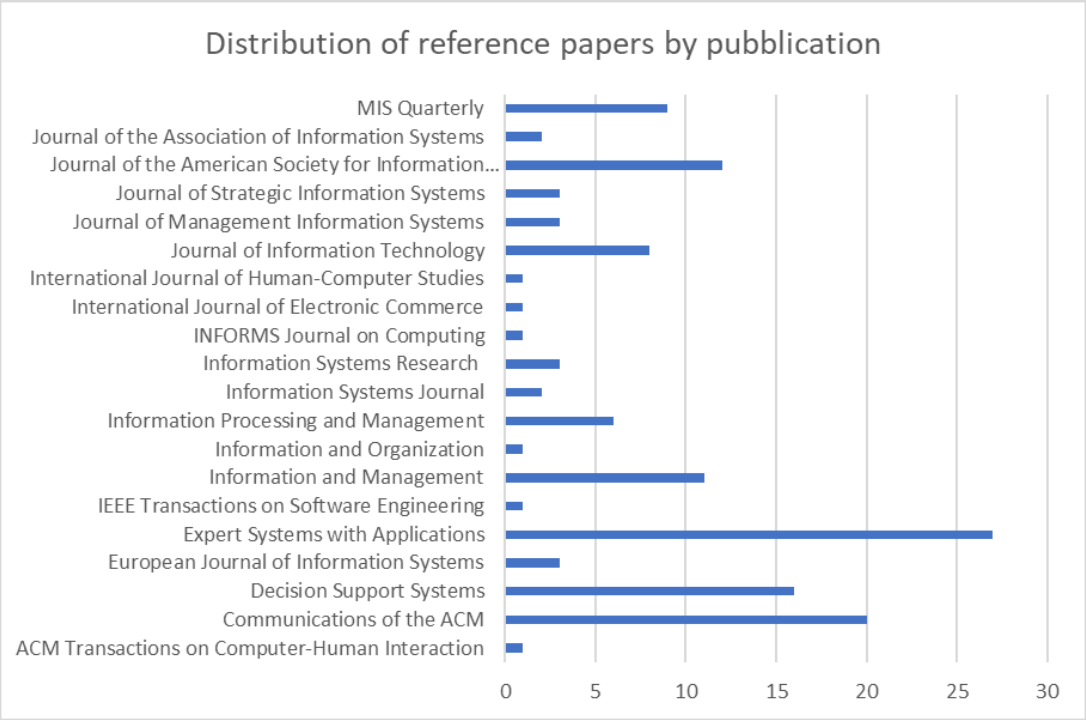

Figure 3: Distribution of paper by journals

As regard the research process **Figure 4**, in the first step 138 articles are found and grouped from 2011 to 2017. The scope of the study is to explore the impact of Big Data on firm's performances and collect all the applications that have been used in different sectors, so it comes that the integration with different platforms and journals excluded from the initial research is necessary. After the second steps 212 references are collected and all the duplicates are then removed. In the last two stages an articles selection in relation to the main topic of the research is performed. Due to the specificity of the argument it is preferred to ignore all the pure computer scientist studies containing line code or different algorithm less useful for the literature review.

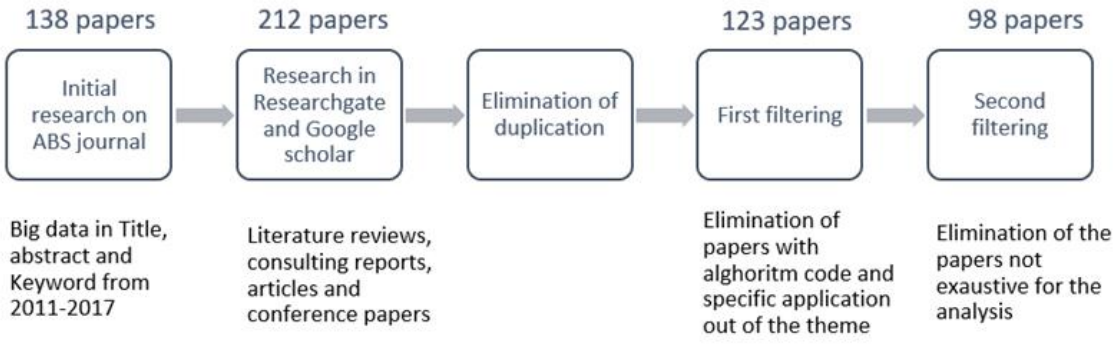

Figure 4: Steps of research methodology

In the end all the articles with specific application not relevant to take out the primary issue related to the research topic are eliminated. The remaining 98 papers are analysed and included in the core collection. According to the defined chapters the articles are then allocated (also more than once) to these last ones to build the main body literature.

Figure 6: Distribution of articles by research type

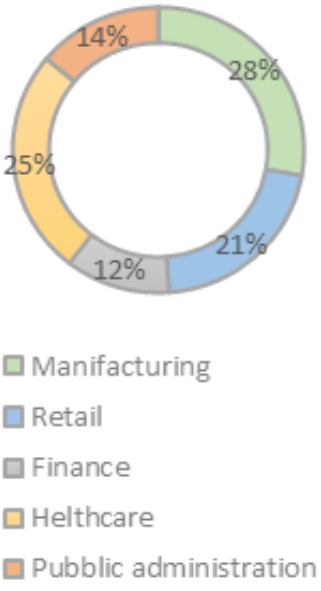

Figure 5: Distribution of articles by research areas

Empirical studies like articles and consulting reports rapresent the 82% of the total papers **Figure 5**, improving and adding specific value for the realization of quantitative frameworks, solutions and alternatives to the big data issues. The other 18% is focalised on the extending of conceptual literature. In the second chart are shown the different areas in wich the articles are allocated and developped. The scholars effort are concerded on the understanding of application in healthcare and manufacturing sectors with a 53% of the total literature. The other 47% is splitted by retail, pubblic administration and finance industries **Figure 6**.

The limitations of the research may be mentioned and considered to underline the variety of this topic. In this case several implications are possible: different interpretations, meanings, and approaches customized on the specific research field.

4. Big data impact on firm's performances

The aim of this chapter is to develop a framework of big data capabilities establishing the impact on firms' performances. In the next sections an analysis of several studies focused on quantitative model about the existent relation between analytic capabilities and possible outcomes is conducted.

4.1. Theoretical assumptions

The big data capabilities research model, with big data components and big data impacts on firm performances (FPER) implies recall of different strategic concepts. The literatures, focused in mapping the quantitative/qualitative effects of using large datasets to support decision (McAfee & Brynjolfsson, 2012), draw their analysis according to two main theories: Resources Based Theory (RBV), and Dynamic Capabilities (DC). In the last two decades these philosophies are gaining importance across scientific literature, useful in different field of application, from strategic reports to technological ones.

4.1.1. Resourced Based View

Resources Based Theory relies on the assumptions about how firms' resources-based can perform better than others in the same industry, how they can gain more competitive advantage. First, even if they operate in the same market there is a usage of various mixture of resources (Peteraf, 2003). This assumption implies that firms can manage different assets that brought value added on their business. Second, the exchange of benefits across resources is difficult, so the capability to fit assets allow sustainable competition over time (Akter S., 2016). In support to Resources Based View (RBV) the VRIO analysis (**Figure 7**) states that firm's resources and capabilities must be valuable (V), rare (R), inimitable (I) and properly organized (O) to ensure FPER qualities. The VRIO Analysis is developed by Barney in the 1991 and is to evaluate the Financial resources, Human Resources, Material resources, Non-material resources (information, knowledge) of the organization. The valuable attribute is referring to resources and capabilities economic value, and in the strategic terms if

they can neutralize threats or to seize an opportunity. Second, the rare dimension indicates if the resource is commonplace or rare among the competitors. The imitable dimension implies if the competitor can obtain and at what cost the resources. At the end of the process, cultures and organization must support the development of VRI resources (Barney J. B. & Hesterly S. W., 2011).

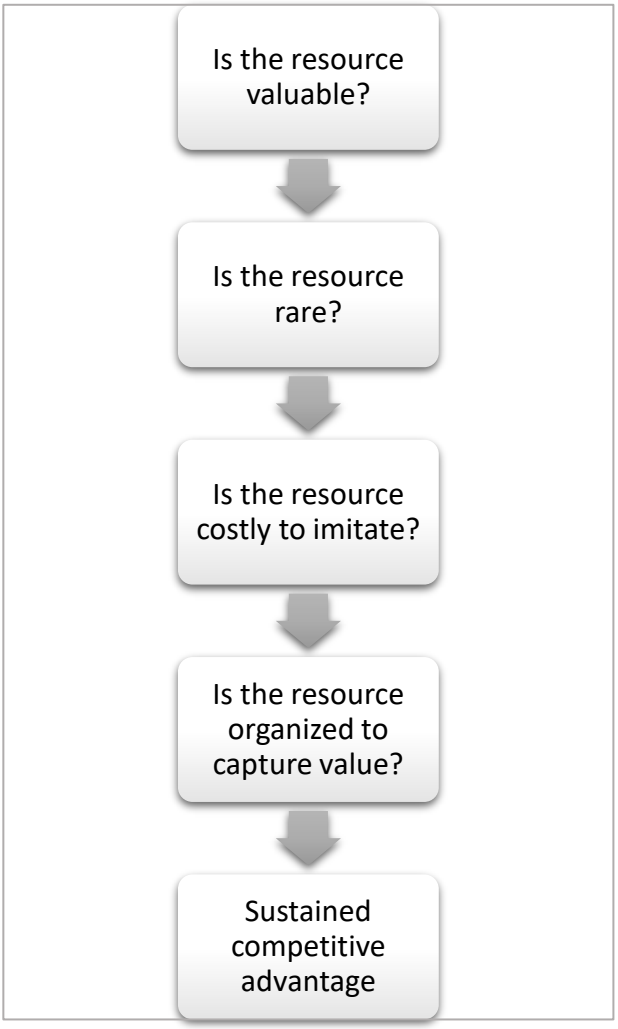

Figure 7: VRIO process

But what are the resources? The most common classification divides the resources in three categories: tangible, intangible and human resources. The first group is referred to physical assets, the real benefits own by companies. Machinery, infrastructure, software and data warehouse are some examples. The intangible is everything regarding not physical assets as brand reputation, organizational culture, patents, etc.

At the end human resources are much more skill than resources. Training, experience, loyalty belong to this group (Grant R. M., 1991).

4.1.2. Dynamic capabilities

Dynamic capabilities is a concept used by several authors to describe how firms integrate, build, and reconfigure internal and external competences to address rapidly and unpredictably changing environments as markets or industries (Teece D. J., 1997; Helfat C.E. & Peteraf M.A, 2003). According to the previous definition some authors argue that agility is the main dynamic capability or higher order dynamic capability, that allows a rapid identification and response to an unexpected shifting in the current business. Different utilization of the resources, reallocation of the responsibilities and share value within the organization are the typical responses of a dynamic company.

But not all the theory used by the authors to develop the framework in the next chapter are cited. Several studies in support to Resources Based View and Dynamic Capabilities have utilized other assumptions to reinforce their researches. Sociomaterialism theory (Akter S., 2016) and Knowledge Based View (KBV) (Côrte-Real N., 2017) are part of them. Knowledge Based View is connected to Dynamic Capabilities because Dynamic Capabilities theory can solve the problem of sustaining competitive advantage in turbulent environments and Knowledge Based View explores a firm's potential to acquire competitiveness in a dynamic market context (Grant R. M., 1991).

Sociomaterialism is referred to the integration of social and material. It is built on the integration between technology, work and organization. The consequences or the main principles is that the resources are so connected that is difficult measure their single contributes (Orlikowski W. J., 2007)

4.2. Big data components

The first part of the framework in **Figure 10** represents the roots of the three levels diagram: big data components (BDC). All the promises that big data brings, need a solid support to transform potential benefits in real improvements. Drawing on Resources Based View the components of big data capabilities are classified in three macro dimensions: tangible, intangible and human resources. The researches collected in **Table 4** agree on the assignments of seven indispensable assets owned by companies to succeed in this technology.

The resources components of BDC are in detail: data, basic resources, technology, technical skills, managerial skills, organizational learning, data driven culture. (Gupta M., 2016). In the further chapters a more detailed description is provided.

4.2.1. Tangible resources

The tangible resources are real assets that firms use to operate and that can be bought or sold in markets. The research model individualizes as the first tangible resources the data. According to Mckinsey report, data is considered an important production factors (Manyika, 2011). Data are now created by multiple sources and in IT sector are growing in Terabytes per seconds. RFID, social networks, sensors, GPS applications and log-files are only few of the variety from data arise down. Information can belong both from internal and external sources (Zhao J.L., 2014). Internal sources create data from firm's internal operation such as inventory updates, accounting transactions, sales, and human resources management (Gupta M., 2016). Common pattern among companies is a data driven decision-making culture according to the internal data (McAfee & Brynjolfsson, 2012). On the other hand, sources as web, GPS locations, mobile phones, cookies and e-commerce communities can be classified as external data. They can provide new insights value in comparison to internal data and can also provide discovery of common patterns in consumer preferences. Firms interested in creating BDA capabilities must integrate their internal and external data (Kwona O., 2014).

The second tangible resource individualized by literature and included in the framework is technology. Data are fast, structured and unstructured so specific tools

to handle this large amount of information became necessary. Firms need to establish a platform that is capable of storing and analysing large amounts (volume) of info continuously flowing in real time (velocity) from many different sources (variety) (Davenport, 2012). The traditional relational database, sequential programming language and legacy system are unable to perform this new event. So, technology as Hadoop, non-relational database (NoSQL), dynamic program language (Python, etc..) allows firms to deliver analytical results in accessible and understandable way in order to support business decision-making (Chen, 2012). These tangible resources can be grouped in three macro categories: big data analysis platforms and tools, databases/data warehouses, programming Languages. They are subdivided according to process flow on information **Figure 8**: (i) Acquisition and recording, (ii) extraction and cleaning, (iii) integration, aggregation and representation, (iv) analysis, visualization and interpretation. (Gandomi A., 2015)

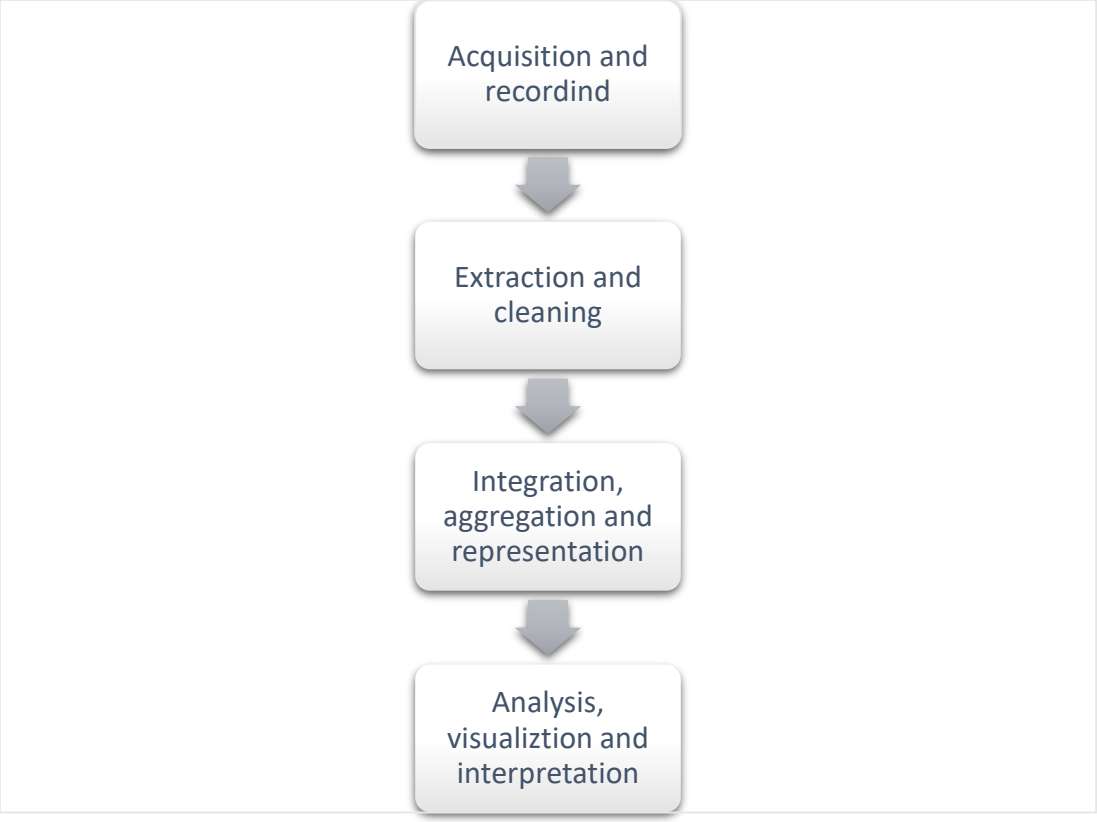

Figure 8: Processing data

The storage of huge amount of data requires new technique and paradigm. The recent trends shift the use of SQL in "Not only SQL" databases, as Cassandra, Mongo that don't use mathematical concept of relations to store the data but have no relation algorithm to achieve higher flexibility. The main advantage is the improvement of the scalability that allows lower investment in hardware. There are four types of database: (I) columnar database, (II) documents database, (III) graphs database and (IV) key-value database. The main difference is in the structure, a graphs database is more capable than others to record graphs and so on.

For the second and third steps open software from Apache are the main actors: Hadoop and algorithm MapReduce. Hadoop is a scalable fault-tolerant distributed system for data storage and processing that is often integrated with MapReduce, a simple but powerful model for large-scale computing that use large number of clusters of commodity PCs (32 gigabytes of memory and1 of RAM) to achieve automatic parallel processing and distribution. (Addo-Tenkorang R., 2016). R, Python are also open sources program language useful to perform more statistical analysis on what data have insights.

And in the end, but not less important there are all the visualization tools to better display the results or the analysis like QlikView, Qlik Sense and all graphs and cluster technology. These technologies take interactive data, and with simplicity and velocity, elaborate a business data model grouping data in graphs and tables dynamically accessible allowing a more intuitive scheme in support to manager decision.

Between data and technology, a reasonable investment decision is fundamental. Firms need to invest in tools and must understand the procedures related to the operations. It requires money but mainly time to reshape business, capabilities and reach the Break Even Point. For this reason, the third tangible resource that is included in the framework is basic resource that include two essential objects: time and investments (Wixom B. H., 2007).

4.2.2. Intangible resources

According to RBV the intangible resources are asset not documented on financial statements but not for this with less value (Grant R. M., 1991). In information technology sector, intangible resources can be source of relevant competitive advantage. In the framework are individualized two main not physical assets: organizational learning and data driven culture (Gupta M., 2016).

The reasons why big data projects sometimes result unproductive is related to organizational culture rather than to the inconsistency of data and lack of technology. Organizational learning is seen as a prerequisite of dynamic capabilities. Referring to KBV the possibilities to transfer old and new knowledge within the organization chain is a fundamental resource to sustain business performance. So, it is related to an organization's developing and create new thinking to enhance existing resources, it is a process through which firms explore, store, share, and apply knowledge. Coordination and control, two main steps of organizational learning, ensure cross-functional synchronization of analytics activities across the firm (Akter S., 2016). The hypothesis is that companies with effective organizational processes in these areas will choose new business intelligence systems and will be able to integrate and manage the information arising from the new tools (Chasalow, 2015).

The second resource is data driven culture, an expression used to identify the firm's capacity to create initiatives related to information inside data. A key success factor to realize the full potential big data project is the application of intuition and experience of top executives with big data analyst results. Most companies still to do initiatives relying only on decisions of important figures which are commonly individualized as the highest paid person's (McAfee & Brynjolfsson, 2012). All employees are required to make some decisions, so it is pertinent to diffuse the culture of data-driven decision-making (DDDM) to every level of organization and regardless of job titles. Both big data physical resources and big data human skill can be bought, instead acquire organizational learning and DDDM capabilities is an intern process hard to copy or imitate (Meyer-Waarden Lars, 2014).

4.2.3. Human resources

Almost all the technology tools about big data are open. Database, computational programs are now available at market prices. Data are internal to the organization or easily obtainable from a third part. In this context probably, the most important component of big data is the human one. Experience, problem solving, leadership, learning time and knowledge are typical characteristic of human resources. These can be acquired from external source hiring people, or internally creating formation courses for specific areas. According to the literature the two identified human components of big data capabilities are: technical skill and managerial skill (Gupta M., 2016). The IT companies' departments count business analysts, data warehouse managers, but especially data scientists whose main technical abilities are in programming, statistic languages, mathematics and business knowledge. They are able to find patterns in large quantities of multistructured data and transform into useful and actionable insight (Davenport, 2012). These results are necessary for top management department that can support their decision with empirical valuation. The new value added of big data is the empiricism hidden in the new way to support the initiatives.

Managerial skills have to be considered as another distinctive component of big data capabilities and are more difficult to find in the market. Decisions born from data have a scientific base but need to be supported by manager intuition and experience. Another skill of manager is the internal capability to across or reduce walls between organizational levels to better perform decisions and information flows.

So, the first level of the graph in **Figure 10** represents the antecedents of the three big data capabilities. The relation between resources and antecedents are tested by different scientific researchers and have the presumption to cover a wide area. However more studies should be integrated with the research model for a more accurate description of the prerequisites to better manage big data technology.

Table 4: Big data resources and references

BD components	BD resources	References
Data (internal, external)	Tangible resource	Gupta et al. (2016), Akter et al. (2016), L. C. Chasalow et al. (2015), O. Kwon et al. (2014), R. Vidgen et al. (2017), R. Cosic et al (2012), Ren et al. (2016), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), Demirkan & Delen (2013), S. Erevelles et al. (2015), G. Shanks et al (2011), L. Kung et al. (2015), R. Dubey et al. (2017), A. Gunasekaran et al. (2017), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)
Technology (Hadoop, NoSql.)	Tangible resource	Gupta et al. (2016), Akter et al. (2016), L. C. Chasalow et al. (2015), R. Vidgen et al. (2017), R. Cosic et al (2012), Ren et al. (2016), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), Demirkan & Delen (2013), S. Erevelles et al. (2015), G. Shanks et al (2011), L. Kung et al. (2015), R. Dubey et al. (2017), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)
Basic resource (time, investment)	Tangible resource	Gupta et al. (2016), Akter et al. (2016), R. Cosic et al (2012), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)
Technical skill	Human resource	Gupta et al. (2016), Akter et al. (2016), L. C. Chasalow et al. (2015), R. Vidgen et al. (2017), R. Cosic et al (2012), Ren et al. (2016), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), S. Erevelles et al. (2015), G. Shanks et al (2011), L. Kung et al. (2015), R. Dubey et al. (2017), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)
Managerial skill	Human resource	Gupta et al. (2016), Akter et al. (2016), L. C. Chasalow et al. (2015), R. Vidgen et al. (2017), R. Cosic et al (2012), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), S. Erevelles et al. (2015), G. Shanks et al (2011), L. Kung et al. (2015), R. Dubey et al. (2017), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)
Organizational learning *	Intangible resource	Gupta et al. (2016), Akter et al. (2016), L. C. Chasalow et al. (2015), R. Vidgen et al. (2017), R. Cosic et al (2012), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), S. Erevelles et al. (2015), L. Kung et al. (2015), R. Dubey et al. (2017), A. Gunasekaran et al. (2017), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)
Data driven culture	Intangible resource	G. Shanks et al (2011), O. Kwon et al. (2014), Gupta et al. (2016), Akter et al. (2016), L. C. Chasalow et al. (2015), R. Vidgen et al. (2017), R. Cosic et al (2012), G. Hassna et al. (2016), Lars Meyer-Waarden et al. (2016), S. Erevelles et al. (2015), L. Kung et al. (2015), R. Dubey et al. (2017), A. Gunasekaran et al. (2017), Chen et al. (2015), P. Mikalef et al. (2017), S. F. Wamba et al. (2017)

Note: Some articles are counted more than once because they cover more than one components

4.3. Big data capabilities and performances

"Big data capability is broadly defined as the competence to provide business insights using data management, technology and talent (human capability) to ensure value creation and competitive advantage" (Kiron, 2014). Using big data, firms potentially are allowed to convert information from data into intuitions and intelligence delivered when and where they are needed (Chen D. Q., 2015). Drawing on the literature on Resources Based View, Dynamic Capabilities, Sociomaterialism and business knowledge, this study puts forward the research model in **Figure 10**. In this model, big data capabilities have significant impact on strategic, financial, market and operational performance, which in turn will influence the firm performances. The analysis is articulated in different steps. First of all, the paper with quantitative and qualitative valuation of relation between BDC and FPER are taken into account. Subsequently are individualized the sub components of firm performance descripted in the four macro categories previously mentioned. The literature research conducted through statistical analysis, Delphi studies, web surveys supported by mathematic relations or real cases study have been the sources. The research model drawn on previous model of Chen et al. (2015), Ren et al. (2016), Akter et al. (2016) Gupta et al. (2016), Gunasekaran et al. (2017), have the aim to extend and clarify existing studies **Table 6**. In the next chapter is performed a description of different big data capabilities, firm performances and then are followed by an extension/integration on the analysis of moderator and mediator direct and indirect effects.

4.3.1. Big data capabilities

To transform resources in potential benefits, companies are involved in routinization, acceptance and assimilation of process (Gunasekaran A., 2017). The three big data capabilities recognized from the literature are Big data management capability (BDMC), Big data technological capability (BDTC) and Big data human capability (BDHC) **Figure 9**. Both capabilities are defined as 'the ability to utilize resources to perform a business analytics task, based on the interaction between the resources (Cosic R., 2012). Business analytics task is referred to use of organizational data, varying from operational activities to management decision-making.

Big data management capability is strictly connected with its antecedent intangible resources and represent the potentiality to ensure a data driven decision making and perform an IT organizational learning. The technical challenges of using big data are a real open point. But the managerial challenges are even greater starting with the role of the senior executive team (McAfee & Brynjolfsson, 2012). Big data technological capability is attributed to the capacity to integrate and manage proper of firm whose assets are composed by big data software, architectures and data. Although attention to technology isn't sufficient, it is necessary component of a big data strategy (McAfee & Brynjolfsson, 2012). Big data human capabilities belong to technical skills of data analyst, data warehouse manager and strategic department. These people typically have also other qualifications in business, management, production process and information technology. Researcher agree that the unique combination of all three capabilities generates a firm specific BDA capability (Gupta M., 2016). The intent of this literature is collect all the researches that describe how the capabilities, with the roots on resources, have impact on the four different firm performances.

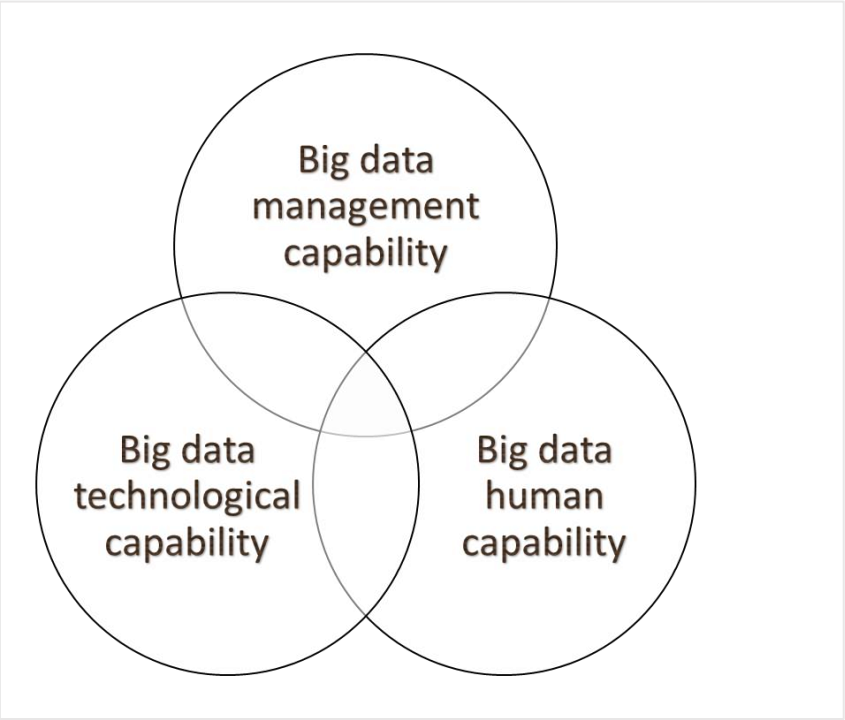

Figure 9: Big data components

4.3.2. Strategic performances

The first relation between dependent variable and independent variable in framework regards strategic decisions and big data tools.

The definition of strategic performances is developed through four main statements: building for business innovations, alignment of resources and planning, building cost leadership, generating product differentiation (Sharma R., 2014; Erevelles S., 2016; Côrte-Real N., 2017; Kache F., 2017; Wamba S. F. F. R., 2017) **Table 5**. The changing in focus toward a network structure of companies is affecting the management of innovation and product design, which is increasingly moving away from a pure internal focus to an external environment understanding and baseline. In particular although companies are still analysing their internal business to improve efficiency, there are many decisions and programs generating from external factors inputs. The Big Data Analytics is moving in this direction, enhancing the innovation and product design. New insights gained from data can be utilized to improve the competitive position of the firm. Strategic decisions based on existing business models can be fine-tuned through more granular data sets. In addition, knowledge gained through increased availability of information is useful to think and develop and reach new business opportunity (Kache F., 2017). For example, the sale of "excess" information goods with little value to companies for which the information holds high value may be exploitation of information through a new corporate business model. The last two strategic performances are main actor of strategy literature, with several examples of how they are adopted in different sectors (Porter M. E., 1985). Cost leadership and differentiation are often sons of separate way of thinking. For instance, Ryanair is focused on cutting cost campaign to reach competitive advantage opposite to Prada or famous high fashion brands that point on bigger willingness to pay of a segment consumer realizing different goods or services. But in both to cases Business analytics enhance to add more value identifying opportunities to improve efficiency (saving cost) (Davenport T., 2014) and customizing products for more appeal. In this analysis the decision of separating strategy and market sphere is due to necessity to associate the decision taken internally considering external environment with performances that see companies comparing each other's.

4.3.3. Market performances

Isolating mediator and moderator aspects, several studies are conducted to extrapolate the relation between big data technology and market. The literature highlights a positive impact of BD capabilities on market performances **Table 5**. Companies, public organizations and in general people are trying to better perform the current activities. Old and current research point out on the understanding of what is the firms' behaviour in the market, what is the growth rate and the market share depending on sectors. For this purpose, three statements are selected to better explain what improvement of market performance means : "Reach new markets more quickly than competitors", "Introduction of new products or services into the market faster than competitors", "Market share bigger than competitors" (Shanks, 2011; Kung L., 2015; Gupta M., 2016; Akter S., 2016; Meyer-Waarden Lars, 2014; Wamba S. F. F. R., 2017). The shared objective of big data application is reduction of the time to carry out a process decision. The rapidity is often translated in profitability and revenues for firms. In the era of Information Technology, lots sectors are losing boundaries and new markets are growing and becoming interesting. With analysis of information it would be possible to find specific customer trend that brings you in new market where operate. Another important aspect of this technology is the transparency of information possible in the different activities, with consequently decrease of risk for business decision and investment opportunities. Companies are installing price determination tools to follow the consumer satisfaction and maximize profitably (Davenport T., 2014). NBA tickets prices change every minute, from sectors to other allowing full load factors and high margin. In the Web 2.0 environment, big data is expected to play a role in global business (Jie Sheng, 2017). Online environment scanning using Web 2.0 helps improve decision making on cross-border merger and acquisition and enforce market share (Lau, 2012). These three level should be examined with the same weights and attention to ensure continuous sustainability and progress.

4.3.4. Financial performances

Companies major worry is the financial integrity of assets and business. But to clarify this issue, it is chosen to consider only validated parameters. In this chapter a description of what are the key performances indicator used from the literature, to define financial performances, is provided. Erik Brynjolfsson (2016), Gupta et al (2016), Akter et al. (2016), Corte real (2017), Gunasekaran et al. (2017), have identified positive effect of big data technological capability, big data management capability and big data human capability on ROI, ROS and Cash Flows **Table 5**. The relation exploded underlining a strictly connection between financial and operational performance. If there is an increasing of operation efficiency a firm improves also the cash flows due to the saving costs and better performance. However, the scope of the research model is referred to isolation of capabilities effect on different business units. Corte real (2017) in his work collected response of an interview, conducted among IT companies, on the improvement of Financial performance due to Big data technology. The three statements with major audit are: "Our EBIT (earnings before interest and taxes) is continuously above industry average"; "Our ROI (return on investment) is continuously above industry average"; "Our ROS (return on sales) is continuously above industry average". To support the relation other authors, describe real cases on companies' big data initiatives. Intesa Sanpaolo Spa dedicated an entire IT department on Big data programs whose main scope is clustering the customers/consumers analysing logfiles in apps, transaction codes in order to better perform loan and specific credit line, consequently with an increasing of the profitability. Furthermore, academic research identify how market react in relation to investments in big data technology such as in Hadoop system (Tambe P., 2014). In the chapter 7 are then described real implication on financial sector by big data.

4.3.5. Operational performances

Operational performances are giving the major attention from the literature. In the research framework in **Table 6** almost 36% of the total references are trying to connect BD capabilities to improvement of Supply chain efficiency, one of the most common goals for any industry (Nguyen T., 2017). In particular the operational performances components are: Customer service improvement, Reducing inventory cost, Asset productivity (ROA, Turnover), Cycle time reduction, Quality improvement (Manyika, 2011; McAfee & Brynjolfsson, 2012; Chasalow, 2015; Chen D. Q., 2015; Gunasekaran A., 2017; Vidgen, 2017; Bughin, 2016; Kache F., 2017; Singh A., 2016; Wamba S. Fosso, 2017). The customer loyalty is particularly easy to lose but it is highly valuable for the organization. In a competitive market, know which offer or service fit for a specific customer, could increase competitive advantage. This is where big data comes in. Collecting data about the buying habits, delivery choices, payment methods, shopping seasons of the customer would give the organization proper knowledge about the consume and pitching the right offers and services (Vidgen, 2017). In a Supply chain perspective, the purpose of inventory is to act as a buffer in stop production cases as breakdown of machinery, unavailability of raw material, semi-component and human resources. Using big data, it is possible to reduce the level of inventory. For example, analysing performance data of the machines can help in formulating preventive algorithms to service the machines before they break down (Singh A., 2016). In this way companies as Pirelli are implementing big data application mixing python language and sensors on machine to reduce production breakdown.

Current researches have discussed the incorporation of social media and web data in operation process and production to achieve a higher asset productivity (Chen D. Q., 2015). With collection of data it is possible to better perform the production step allocating the right flow of goods optimizing the process. Big data allow firms to connect systems with millions of customers. This would allow them to understand behaviour and match their preferences. Next step could be trying to move goods and services reducing lead time. For example, delivery company as Amazon, DHL, could deliver directly to a warehouse near the intended consumer address and shipped out to their destination the next day without any delay. Big data could raise the quality of

information and data, and according to the extant big data literature, the framework identifies the importance of technology and information quality on critical operational outcomes as success factors of BDA projects (Ren S. Ji-fan, 2016). As example of what improvement of performance means it represent by ORION project (On-Road Integrated Optimization and Navigation) from UPS which saved more than 8.4 million gallons of fuel by cutting 85 million miles off of daily routes. The total program led to company $30 million of save cost (Ziora, 2015).

Table 5: Firm performances

Firm performance	**Key performance indicators and components of FPER**
Financial performance	• Cash flows • Ros • Roi

Operational performance	• Customer service improvement
	• Reducing inventory cost
	• Asset productivity
	  ○ ROA
	  ○ Turnover
	• Cycle time reduction
	• Quality improvement

Strategic performance	• Alignment of resources and planning • Building for business innovations • Building cost leadership • Generating product differentiation
Market performance	• Reach new markets more quickly than competitors. • Introduction of new products or services into the market faster than competitors. • Market share bigger than competitors.

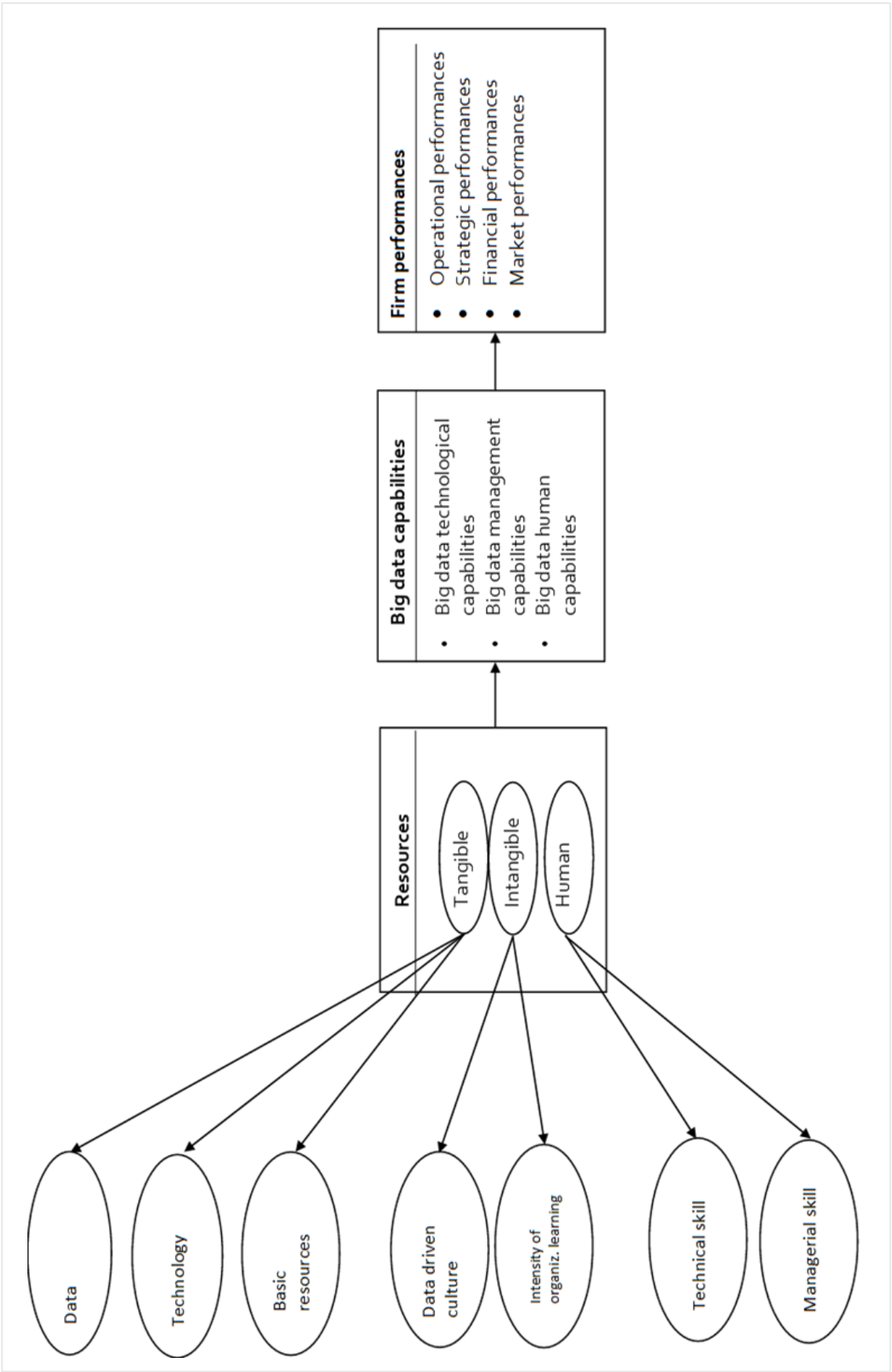

Figure 10: **General Big Data model**

Table 6: Big data capabilities and Firm Performances

Performances	Sub category of BD Capabilities	References	#	%
Market performances	big data technological capability	G. Shanks et al (2011), L. Kung et al. (2015), Garmaki M. et al (2016), Gupta et al (2016), Akter et al. (2016), Lars Meyer-Waarden et al. (2016), G. Hassna et al. (2016), Ren et al. (2016), P. Mikalef et al. (2017), S. F. Wamba et al. (2017).	10	5,7%
	big data management capability	L. Kung et al. (2015), Garmaki M. et al (2016), Gupta et al (2016), Akter et al. (2016), Lars Meyer-Waarden et al. (2016), G. Hassna et al. (2016), P. Mikalef et al. (2017), S. F. Wamba et al. (2017).	8	4,5%
	Big data human capability	G. Shanks et al (2011), L. Kung et al. (2015), Garmaki M. et al (2016), Gupta et al (2016), Akter et al. (2016), Lars Meyer-Waarden et al. (2016), P. Mikalef et al. (2017), S. F. Wamba et al. (2017).	8	4,5%
Operational performances	big data technological capability	G. Shanks et al (2011), McAfee and Erik Brynjolfsson (2012), P. Tambe (2014), L. C. Chasalow et al. (2015), Chen et al. (2015), Kung et al. (2015), Ghazwan Hassna et al. (2016), Garmaki M. et al (2016), Gupta et al. (2016), Akter et al. (2016), Erik Brynjolfsson (2016), Corte real et al. (2017), A. Gunasekaran et al. (2017), Richard Vidgen et al. (2017), Jacques Bughin (2016), F. Kache et al. (2017), Ankit Singh et al. (2017), S. F. Wamba et al. (2017).	22	13,4%
	big data management capability	McAfee and Erik Brynjolfsson (2012), L. C. Chasalow et al. (2015), Chen et al. (2015), LeeAnn Kung et al. (2015), L. C. Chasalow et al. (2015) Ghazwan Hassna et al. (2016), Garmaki M. et al (2016), Gupta et al. (2016), Akter et al. (2016), Corte real et al. (2017), A. Gunasekaran et al. (2017), Richard Vidgen et al. (2017), F. Kache et al. (2017), Ankit Singh et al. (2017), S. F. Wamba et al. (2017).	18	10,9%
	Big data human capability	G. Shanks et al (2011), McAfee and Erik Brynjolfsson (2012), P. Tambe (2014), Chen etal. (2015), LeeAnn Kung (2015), Jacques Bughin (2016), Erik Brynjolfsson (2016), Garmaki M. et al (2016), Gupta et al (2016), Akter et al. (2016), Corte real (2017), A. Gunasekaran et al. (2017), Richard Vidgen (2017), F. Kache et al. (2017), Ankit Singh et al. (2017), S. F. Wamba et al. (2017).	19	11,52%
financial performance	big data technological capability	G. Shanks et al (2011), McAfee and Erik Brynjolfsson (2012), L. C. Chasalow et al. (2015), Chen et al. (2015), LeeAnn Kung et al. (2015), Lars Meyer-Waarden et al. (2016), Garmaki M. et al (2016), Ghazwan Hassna et al. (2016), Jacques Bughin (2016), Ren et al. (2016), Akter et al. (2016), Corte real et al. (2017), A. Gunasekaran et al. (2017), F. Kache et al. (2017), Ankit Singh et al. (2017), P. Mikalef et al. (2017), S. F. Wamba et al. (2017).	17	10,25%
	big data management capability	McAfee and Erik Brynjolfsson (2012), L. C. Chasalow et al. (2015), Chen et al. (2015), LeeAnn Kung et al. (2015), Akter et al. (2016), Lars Meyer-Waarden et al. (2016), Garmaki M. et al (2016), Ghazwan Hassna et al. (2016), Corte real et al. (2017), A. Gunasekaran et al. (2017), F. Kache et al. (2017), Ankit Singh et al. (2017), P. Mikalef et al. (2017), S. F. Wamba et al. (2017).	14	8,18%
	Big data human capability	G. Shanks et al. (2011), McAfee and Erik Brynjolfsson (2012), Chen et al. (2015), LeeAnn Kung et al. (2015), Lars M. Waarden et al. (2016), Jacques Bughin (2016), Akter et al. (2016), Garmaki M. et al (2016), Corte real et al. (2017), A. Gunasekaran et al. (2017), F. Kache et al. (2017), Ankit Singh et al. (2017), P. Mikalef et al. (2017), S. F. Wamba et al. (2017).	14	8,33%
Strategic performance	big data technological capability	G. Shanks et al (2011), R. Cosic et al (2012), R. Sharma* et al. (2014), S. Erevelles* et al. (2015), LeeAnn Kung et al. (2015), Akter et al. (2016), L. M. Waarden et al. (2016), Ghazwan Hassna et al. (2016), Corte real et al. (2017), F. Kache et al. (2017), S. F. Wamba et al. (2017).	13	8,33%
	big data management capability	R. Cosic et al (2012), R. Sharma* et al. (2014), LeeAnn Kung et al. (2015), S. Erevelles* et al. (2015), L. M. Waarden et al. (2016), Ghazwan Hassna et al. (2016), Akter et al. (2016), Corte real et al. (2017), F. Kache et al. (2017), S. F. Wamba et al. (2017).	11	7,05%
	Big data human capability	G. Shanks et al (2011), Cosic et al (2012), R. R. Sharma* et al. (2014), S. Erevelles* et al. (2015), LeeAnn Kung et al. (2015), Akter et al. (2016), L. M. Waarden et al. (2016), Corte real et al. (2017), F. Kache et al. (2017), S. F. Wamba et al. (2017).	11	7,05%
			156	100%

Note: Some articles are counted more than once because they cover more than one type of business value
* The relation between BD capabilities and performances is not supported

4.4. Integrative research model

The model presented covers a wide range of analysis, from surveys research to statistical categorization. The results show the positive effect of big data capabilities on firm's performances. But some researchers have investigated only the direct relationships and ignore mediating effects completely. This focus can change the interpretation of the results when a variable has no direct effect because its effect is mediated by another variable. And in a worst case that variable is considered not relevant for answering to the question at all. (Hair, Ringle, & Sarstedt, 2013). Indirect or mediated effects constitute a type of relationship between constructs that often occurs in partial least squares path modelling (PLS) (Nitzl C., 2016). The procedure for mediation analysis is based on the coefficients and standard errors of the direct paths between independent and mediating/moderating variables and mediating/moderating and dependent variables (Wamba S. F. F. R., 2017). To better create the big data framework Akter et al. (2016), P. Tambe (2014), Chen et al. (2015), R. Dubey et al. (2017), performed an integration of the relation BDC and FPER with the effects of mediators and moderators, this integration and analysis is collected to realize the framework in **Figure 11** and **Figure 12**.

4.4.1. Scenario with mediators

In the first scenario are considered as mediator all the capabilities or operations that influence directly the block capabilities-performances **Table 7**. In the multivariance statistic the mediator changes the variables relation, empowering or decreasing the effects. Multiple regression analysis with mediation tests is used to test the hypotheses due to the complexity of the model and to test all the possible implications (Eckstein, 2015). Researchers have developed their model with alternative and integrative hypothesis on agility, top management commitment, marketing capability and business alignment. The biggest influence on Firm Performance is due to the agility, seen as mediator of knowledge assets on performance for all of the three kinds of big data capabilities (Côrte-Real N., 2017). Agility is defined as "capacity of an organization to efficiently and effectively redeploy/redirect its resources to value creating and value protecting (and capturing) higher-yield activities as internal and

external circumstances warrant" (Teece D. J., 1997). The mediating role of Top management commitment between resources and big data management capabilities highlights that concrete meta-structuring actions by the top management play a significant role in assimilating big data performances in organizations. Top management needs to be able not only to acquire resources but to commit to this process by orchestrating resources to achieve higher performances. (Gunasekaran A., 2017). Market orientation is referred to all the actions as listening to customers and delivering solutions based on the interests and wants of the customers (Bulent Menguc, 2006). For this reason, it is chosen to cite the third mediator individualized from literature, Marketing capabilities that is recognizing as the firm's ability to understand and meet customer needs and deliver its products and services to customers. This quality reinforces positively the Big Data Capabilities effect especially on strategic and market performances (Meyer-Waarden Lars, 2014). Business alignment derives not from a specific change in organizational structure due to the adaptation, but from the understanding of consumer activities to extract hidden insights. All the previous variables have positive effect on the relation between dependent and independent variables but not all the sides are analysed so it appears to be a fruitful area of research far into the future.

Table 7: Integration with mediators

Mediator of BD capabilities	BD capabilities	Performances	References
Marketing capability	Big data management capability	Firm performances	Sharma et al. (2014), Lars Meyer-Waarden et al. (2016);
Business alignment	Big data management capability Big data technological capability	Firm performances	S. Erevelles et al. (2015), G. Hassna et al. (2016);
Agility	Big data management capability Big data human capability Big data technological capability	Firm performances	LeeA. Kung et al. (2015), S. Erevelles et al. (2015), L. C. G. Hassna et al. (2016).
Top management commitment	Big data management capability	Firm performances	Chen et al. (2015), A. Gunasekaran et al. (2017);

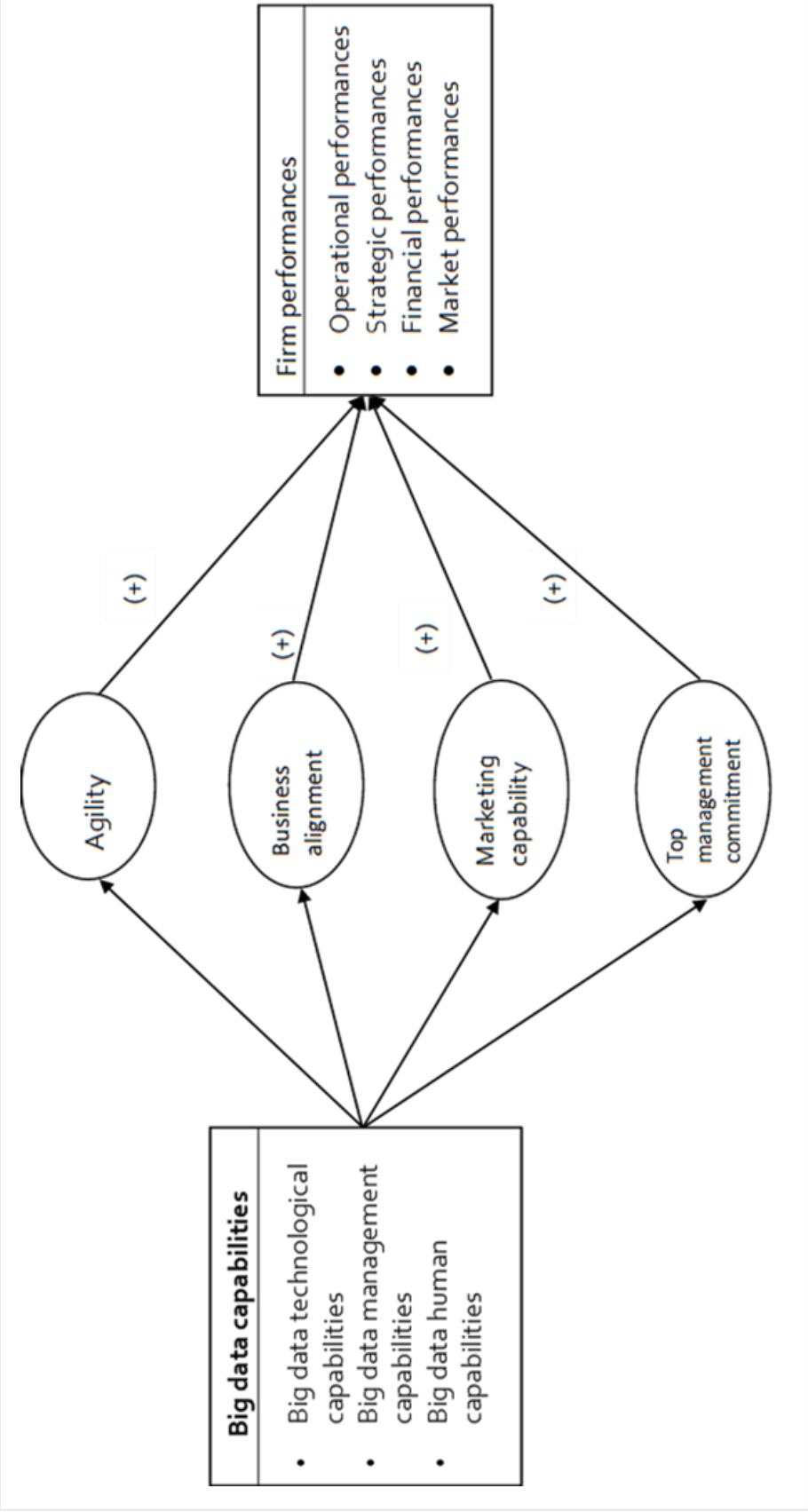

Figure 11: Integrative model with mediators

4.4.2. Scenario with moderators

Talking about direct and indirect effect, moderator variables have the peculiarity to manipulate the relation between Big Data Capabilities and Firm performance. In particular Akter et al. (2016), P. Tambe (2014), (Dubey R., 2017) added to their research model control variables as analytics capability, business alignment and agility to determine the total effect of managing big data resources on strategic, financial, market and operational performance, not ignoring indirect effects **Table 8**. In particular big data management capabilities have strong influence on firm performance in companies able with business alignment and agility, characteristics of a strategic organizational capability that can help firms match resources with changing environmental opportunities. In addition, following this logic it may posit that flexible orientation may strengthen the indirect effect of Big Data Capabilities on Firm Performances (Dubey R., 2017). The benefits of technology and big data programs not always follow the amount of investment in business analytics technology. However, G. Shanks et al (2011), Akter et al. (2016), G. Hassna et al. (2016) have underlined its positive potential indirect effect on operations and financial performances **Figure 12**.

Table 8: Integration with moderators

Moderator of BD capabilities	BD capabilities	Performances	References
Analytics capability	Big data technological capability Big data human capability	Firm performances	G. Shanks et al (2011), Akter et al. (2016),
Business alignment	Big data management capability	Firm performances	P. Tambe (2014), R. Sharma et al. (2014), Akter et al. (2016), Lars Meyer-Waarden et al. (2016)
Agility	Big data management capability Big data human capability Big data technological capability	Firm performances	G. Shanks et al (2011), R. Sharma et al. (2014), Chen et al. (2015); Chasalow et al. (2015), Wamba et al. (2017), Corte real et al. (2017);

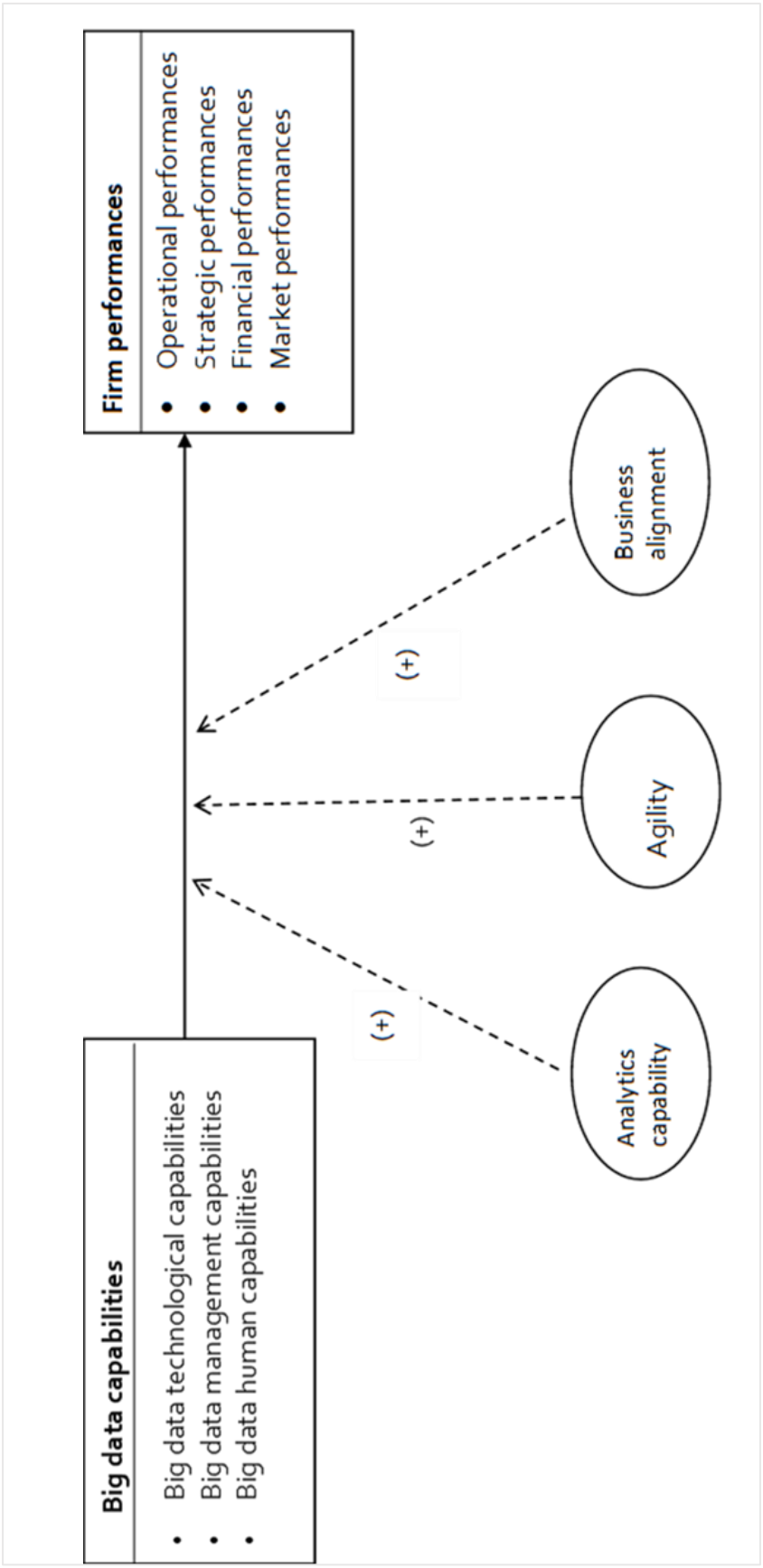

Figure 12: Integrative model with moderators

5. Big data across industries

In recent years there is an explosion of new business, as start-ups, *servitization* (Opresnik D., 2015) and online products, due to the advance of technologies. The boundaries to classify the various sectors are undefined and also the description of their main characteristics because of internet is becoming a hard work for researchers. In this context one of the preeminent innovation are big data that is going to reshape industries. The ability to take advantage of all available information has become a critical ability for organizational success (Olszak, 2016). It is evident that there is a different approach on managing big data. Start-ups, online business are faster and without constrains. The large companies must integrate, merge the old infrastructure and analytics capabilities with the new one. Big data scientist shouldn't be separated from database administrator (Davenport T., 2014).

In the next sections it is first introduced the potential of big data in five domains **Table 10**: healthcare and wellbeing, finance, public administration, manufacturing and retail/e-commerce (Manyika, 2011). The industries configuration allows big data to have a great impact in defining strategic decisions, operational performances and market alignment. First this literature work wants to understand common trends across industries through a presentation of brief case studies and real applications. Second, are defined the sectors in which are applied the principles of the framework previously discussed.

Manyka et al., (2011) in the Mckinsey Global Institute report shows example of the big data impact in term of dollars collecting statistical researches. Focused on five sectors, the most important consequence is in US health care that could save 300$billion per year anticipating problems and supporting the people needs and with a 0,7% per year of productivity growth. As regard Europe public sector administration, 250€ billion are the value per year capable to generate the application of business intelligence with an annual productivity of 0.5%. In the mobile sector, the management of global personal location data could generate 100$ billion of revenues for service provider and up of 700$ billion of value to end users. In the end, as regard companies that provides goods, US retail and manufacturing could increase respectively 60% in

net margin and 50% in development cost of assembly, adding a 15 of annual productivity growth and 7% reduction in working capital.

In **Table 9** there is a gathering of studies grouped by category in which are presented the main topic. The aim of this chart is to give an overview of the different applications customized on different business, exploring the potential of big data in five domains: healthcare and wellbeing, finance, public administration, manufacturing and retail & ecommerce. It also includes a comparison between AS-IS and TO-BE application, programs to emphasize the impact of big data and the understanding of common trends across industries through a presentation of brief case studies and real applications. Therefore, the choice of the articles presented is done according to (i) the level of impacts that the use of big data generates, (ii) the possible solutions' roll out. Further suggestions and studies can implement the list presented below building a common review where find useful guide.

Table 9: Big data and industries

Industry classification	# articles	Description of industries
Healthcare and wellbeing	5	Includes smart technologies that improve the hospital operation and redesign healthcare structure.
Financial	5	How banking and financial organizations manage big data to achieve more competitive advantage; includes also new forms of payment and products.
Public administration	4	Includes the analysis of the relationship between citizen and government in big data era and of the technology tools against terrorism and criminality.
Manufacturing	4	Includes supply chain issues; technologies based on big data to improve operational performance and relation with supplier.
Retail and e-commerce	5	How web-based companies develop big data programs to implement market strategies tailored on customer

Table 10: Applications of big data in the five domains

Sector	Author	Title
Healthcare and wellbeing	(Manyika, 2011)	"Big data: The next frontier for innovation, competition, and productivity"
	(Chen, 2012)	"Business intelligence and analytics: from big data to big impact"
	(Wang, 2016)	"Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations"
	(Krugler, 2016)	"What Happens When Big Data Blunders? "
	(Pramanik Ileas, 2017)	"Smart health: Big data enabled health paradigm within smart cities"
Finance	(Turner D., 2013)	"IBM Analytics: The real-world use of big data in financial services"
	(Gui & Ming, 2015)	"Analysis on Information Value of Big Data in Internet Finance"
	(Chan S. W.K., 2016)	"Sentiment Analysis in Financial Texts"
	(Seddon J.J.M., 2017)	"A model for unpacking big data analytics in high-frequency trading"
Public Administration	(Manyika, 2011)	"Big data: The next frontier for innovation, competition, and productivity"
	(Chen, 2012)	"Business intelligence and analytics: from big data to big impact"
	(Lavertu, 2015)	"We All Need Help: "Big Data" and the Mismeasure of Public Administration"
	(Guillamón M. D., 2016)	"Factors influencing social media use in local governments: The case of Italy and Spain"
	(Anshari M., 2016)	"E-Government with Big Data Enabled through Smartphone for Public Services: Possibilities and Challenges"
Manufacturing	(Manyika, 2011)	"Big data: The next frontier for innovation, competition, and productivity.
	(Opresnik D., 2015)	"The value of Big Data in servitization"
	(Lee J., 2013)	"Recent advances and trends in predictive manufacturing systems in big data environment"
	(Wang G., 2016)	"Big data analytics in logistics and supply chain management: Certain investigations for research and applications"
	(Bumblauskas D., 2017)	"Smart Maintenance Decision Support Systems (SMDSS) based on corporate big data analytics"
Retail and E-Commerce	(Manyika, 2011)	"Big data: The next frontier for innovation, competition, and productivity"
	(Schäfer K., 2013)	"Determining the performance of website-based relationship marketing"
	(Leeflang P. S.H., 2014)	"Challenges and solutions for marketing in a digital era"
	(Lau R Y.K., 2016)	"Editorial: Big Data Commerce"
	(Matthias O., 2017)	"Making sense of Big Data: can it transform operations management?"

5.1. Big data in Healthcare and wellbeing

Wellbeing and health care sector cover a large fraction of GDPI especially in the national economy, facing hard productivity challenge. There are several stakeholders committed in this organization, patients, pharmacist, hospitals, doctors, product providers and payors. For each character there are thousands of data generated with different structure, information and functionalities. However, they are all strictly related and connected. The patient consults the hospital that buys products from pharms provider, all the exchange is realized through payment. Finding technologies that are able to store and analyse in a single silo all these pool of data is attractive but implies also huge investment and time. The structure of this industries is innovative as regard the single machine, or the drugs but the tools to reach and sustain competitive advantage are less investigated. The next cases describe the real state of art of the big data development across different subjects and different application **Table 11**. It can be seen as another starter point to increase the gross demand product index. The social effect of the healthcare financial improvement also depends on the privatization of the organization. There will be also discussion on this new wearable device that raises first the participation of the patients on the prevention despite the solution, and then the quality of the information collected from wellbeing companies.

In the US healthcare industry big data is generating new business value, increasing efficiency and productivity of programs. Manyika et al. (2011) report highlights two To-Be applications or business innovation that compete with existent solutions. The first innovation, regarding the aggregation and the analysis of patient records, consists in providing information and services to external firms. In this perspective US healthcare could support adjacent business with the implementation of large datasets that could help biomarkers in the development of more customized R&D programs and facilities. The second solution is the creation of web platform and communities to share both patient's experiences and to link doctors and wellbeing organization in order to reinforce the information flow.

Chen et al. (2012) emphasize the attention on the shift in the management of the data including different real applications adopted from wellbeing sector. Two are the main sources of the smart health sector: genomics data (DNA, bioinformation) and payer–

provider big data (health records, pharmacy and doctors' prescription, insurance records, patient feedback and reactions). For the first perspective terabyte of data are generated for each person and for payer perspective a matrix with thousands of transactions for each patient must be managed. HER (electronic healthcare records) is a powerful instrument to monitor patient disease widely develop and diffuse among hospitals. With a temporal events relation and doctor diagnosis a better explanation of pattern of the disease progression is performed. Adding clustering, associations rules technique to health care analysis new program as PatientsLikeMe provides more control and support to patients with chronic diseases as Parkinson, diabetes and Alzheimer's. The last example of applications that joins Big data technology and healthcare is the SHB (Smart Health Being) that with sensors, networks and new devices is able to monitor the wellbeing anticipating the diseases.

Y. Wang et al. (2016) focalizes his attention on the issue related to the sustainability of big data programs. In particular in his research draws five strategies to succeed with business analytics. The first strategy is the implementation of a data governance that is useful to harness the amount of information, structured and unstructured, knowledge and array (HER) generating from different applications. The success in management of data governance requires a clear guideline that specify how all the data has to be understood and processed. Then, healthcare organizations must review the data collected and integrate them to success in big data implementation. The second strategy is the development of an information sharing culture, to ensure a right information flow. In particular it is necessary an engagement of data providers from the first stage of the chain to meet standards for data delivery. The third strategy is referred to the training for key figures in the organizations. In managing this technology, it is important to provide analytical courses such as data mining, statistics to ensure the efficiency of the program. The AS – IS model to record data is the even more expensive distributed files system, typical of small-medium size healthcare organization. As in the other sector the To-Be model encouraged from literature is the shift in a cloud computing services, that gives the chance to perform the data stored in a common cloud database. The last strategy suggested from Wang is the generation on new business idea with manager that should support the use of KPIs, reports and interactive charts to discover new trend or develop new ideas.

L. Kugler (2016) put forward the problems or the failure that big data have shown in healthcare organizations. Two major initiatives in this sector were conducted from Google in forecasting the flue of 2013 and Ebola disease in 2014 and in the early 2015. Inaccurate results were recorded, due to a combination of big data inconsistencies and human errors in the interpretation of the information. Researches show how the starting model used rash with preliminary conditions and the random variable of the transmission of the germ was completely ignored.

M. I. Pramanik (2017) analysed the big data technology as input to change the basis from which is founded the society. In his work there is a description of three main characters: city, healthcare and data that are shifted over the years. The first change is from a classical city-health and data to pervasive city-health and data. In the traditional city, the healthcare solution implies that the doctor visit patients with normal technologies and record process into normal data base management system. The pervasive city, have a pervasive healthcare that uses datamining solutions, as clustering, regression to better perform services, to improve operation and wellbeing for people. To do this, the most common goal is to reduce time, decrease barrier and remove location constraints. The last change is from a pervasive city-healthcare and data to a smart city-healthcare and data through big data deployment. Someone argue that digital and smart city, as digital and smart health, are two different ways to describe the same thing. However, the substantial difference is due to the complexity of the technology that support the extremely heterogeneity of the data. A smart city has to manage security, financial constraints and cultural aspect to succeed in this development.

Table 11: Big data in Healthcare sector

Study	Potential research areas	Definition	Purpose
(Manyika, 2011)	Big data in the US healthcare, a quantitative and qualitative analysis.	big data as new way to reduce the gap in US healthcare industry.	The use of BD for cost savings, increase efficiencies, improved treatment effectiveness and productivity.
(Chen, 2012)	Different field of application in healthcare organization	Big data is the support for improve performance	A list of BD initiatives in Healthcare sector is painted.
(Wang, 2016)	Multidimensional benefits framework from BD	BD analytical capability support decision and organization in healthcare.	Identification of five strategies that healthcare sector could use to implement their big data analytics initiatives.
(Krugler, 2016)	Forecast analysis to predict disease	Big data is a tool to predict event associated with the healthcare	The dark side of using BD algorithm in healthcare sector.
(Pramanik Ileas, 2017)	Innovation and implementation of new initiative in the healthcare organization	Big data and smart system are interconnected	Proposal of a big data enabled smart healthcare framework that offers conceptual models of intra and interorganizational business operation.

5.2. Big data in Finance

Banks, Trading associations, funds have as primary asset money and data. The business core is supported by internal data originated from customer transactions, log-files and in general several kinds of payments. These subjects own petabytes of structured and unstructured information. With the advent of big data technologies there is a reshape of the data functionality and values. Innovative banking and financial associations are committed in extract value from uncertain data. In the next lines a collection of what are the main applications, services, programs and major trend are discussed **Table 12.** The literature introduces a small part of the total applications that could be collected and discussed in future work.

For D. Turner et al. (2013) big data is an imperative business challenges not only for banking but for all the financial organization that provide economic solutions and services. The integration of analytics capability with process has impact on the entire industries. The bank peculiarity is the absence of physical products; however, the data became input of information and the most important assets. The main sources are transactions, hundreds of million per day, that fills the database. In this research a survey on bank associations shows how the bigger part of the companies are in the early stages of big data solution. In particular 26% of financial services are focused on the understanding of the big data project, 47% are already organized technology and road map algorithm and only the 27% is ready to launch or have already implemented a big data program. Four are the key finding that are monitor of how banking is managing big data. First the customer analytic is driving the data decision making program, shifting from a product-centric to a customer-centric organization. Second, the major difficulties met from companies is the integration of information delivered by different sources of different type. The creation of a common silos where store and process data are expansive and complex. As in the other industries, also in the financial sector the firm are focusing their effort to manage the internal data, ignoring the data belonging from social media and external source. And in the end big data requires strong capabilities. As already described, the human skills are fundamental antecedents of big data capabilities so the recent trend in banking is to hire also data scientist, engineering to develop the IT department.

Researcher groups have understood the value of big data in the financial organizations with all the risks associated. The banks or payment transaction can't tolerate any security problems that would influence negatively the company business. For this reason, a trusted network trading system, with a software behaviour certification, to build a behavioural certification platform system is developed by Changjun Jiang et al. (2014). This system is able to manage and ensure the security of the transaction allocating to a unique central entity the monitoring of the transaction and of the links between the different departments. Fu Y. Gui et al. (2015) defined the merge between internet spirit and finance industry as internet finance. This union develops a series of services as P2P micro credit, crowd funding, crypto vault, online manage money matters that are the basis of concrete products as Alipay and PayPal. The author conducts a research in which individualizes seven information value that big data generate on internet finance. "Big data can manage the long tail demands of internet finance medium and low-income users effectively" (Day j., 2011). The traditional business of small medium transaction is large in number and with high margin, however the tools are rigid and immobilized. Internet will produce different structure and dynamic information creating the possibility to realize medium and low-income user demand. "Information value analysis on big data in internet finance risk supervision". Often financial transactions are followed by risks of credit insolvency, with big data is possible to monitor different index evaluating the status of the consumer. On the other hand, the consumer could monitor the operation of the internet banking. "Information value analysis on big data in the third-party payment". With the implementation of big data, a third-party payment manages the customer relationship among sellers and buyers, can provide relevant information for other finance services and can mine he customer orientation. "Information value analysis on big data in internet finance micro credit business and crowd funding business". The loan process is complicated due information asymmetries. But the information reached by the use of analytics could improve the micro credit business. In particular big data can find honest and reliable business environment and remove traditional credit boundaries. "Information value analysis on big data in internet finance online manage money matters". Online platforms as eBay and the other relevant operator own massive information that have to be analysed, from their finance department, with big

data technologies in order to guide investment behaviours. And the last value is referred to the improvement of the traceability given by the use of the new technologies: “Information value analysis on big data in internet finance business recommendation” and “Information value analysis on big data in internet finance transaction subject authentication”.

Among the protagonist of financial industry are not only banks but also entire trading market in which small, medium and large size firms operate. Three important activities are included: fund management for investment with long term, High frequency trading by proprietary of financial firms and low frequency trading that represent the traditional broker. Seddon et al. (2016) examines how in the high frequency trading the implementation of big data analytics strategies could generate a competitive advantage on other High frequency trading programs or on the other financial initiatives carried out by low frequency’s. The developed model integrates the 5V’s of big data (variety, velocity, volume, veracity, value) with other two V (visualization and variability) and has the aim to illustrate how HFT can improve performances with this new application. The big data HFT system is capable to collect millions of trades transaction and process real time event to identify LFT activities, giving advantage over the slower competitors.

The texts are mainly source of unstructured data, sentences and word are possible to combine in various and different mix. Chan et al. (2016) realize a work in which develop a SAE (sentiment analysis engine) for financial statements composed by five major modules, namely the (i) feature extraction, (ii) sentence chunkier, (iii) phrase recognizer, (iv) parse tree resolution and (v) sentiment assessment modules. The application shows how the investors could deal with data in a way that eases demands on their cognitive limitations and decrease the negative effects of operating within bounded rationalities.

Table 12: Big data in Financial sector

Study	Potential research areas	Definition	Purpose
(Turner D., 2013)	Business and IT professionals have to work together to take advantage from BD technologies	Big data is the solution to identify business requirements	Banks and financial markets firms must leverage their information assets to gain a comprehensive understanding of markets, customers.
(Gui & Ming, 2015)	Internet finance, an emerging sector which combines traditional finance and Internet.	Big data analysis technology to improve Internet finance transaction efficiency	Improve the information value on big data in internet finance business activities and create value creation
(Seddon J.J.M., 2017)	Potential source of policy maker action	Big data is a source of competitive asymmetries between HFT and LFT market	Give a list of advantages that BD offer to improve HFT performances.
(Chan S. W.K., 2016)	Machine learning tool in finance test	Big data represent a source of information detection	Algorithm and technique to understand pattern in financial field

5.3. Big data in Public Administration

The public sector is always under the lens of the citizens, resulting with high pressure as regard the productivity and the transparency of the social programs. As the Head of State, it is owner of a multitude of sub organs source of enormous pool of data. Given its particular configuration there are several ways in which it could use the information improving community wellness. So general trends are showing interest from Government on big data technologies under the hypothesis to better allocate the resource and to customize action. In addition, the transparency is required from the other size. In the next lines a collection of what are the main applications, services, programs and major trend are discussed **Table 13**.

The philosophy of the State is to provide to the same citizen the same services. Manyika et al. (2011) underlines how segmenting and tailoring services could increase efficiency and people satisfaction. For example, the German labour association analyses its large datasets to find a pattern specific for its worker to better perform "the find a job" service. In this context perform a customized service could increase gross margin productivity and improve commonwealth.

Studies and scientific works in engineering, informatic, computational science, social science, text mining, hardware system and medicine have conducted to help the development of technologies against terrorism, violence, cybersecurity and other threats (Chen, 2012). 33$ billion are the total expense in cyber security for large medium size companies. As one of the main Public Administration tasks there is the public safety. Intelligence and Security department are gathering structured and unstructured datasets from different sources as criminal records and people charges. However, some applications to manage these data lack in technology and framework. Nowadays business intelligence with network analysis, rule mining and spatial temporal classification should be considered for security informatics development.

An interesting work from Guillamón et al. (2016), whose focus is to analyse the interaction between citizens and public sector, is depicted to understand how government could use big data application to improve communication. The importance of transparency is declared across industries, from manufacturing to healthcare. In particular, as regard the policy field, it is an essential tool to provide better knowledge

of state programmes and to increase the trust in the public sphere. The big data enhance an ex-ante political control, a day by day responses and an ex-post monitoring of the initiatives. The research collect data from different Italian and Spanish cities in which government is utilising social media to diffuse programs and to communicate with citizens. In this way the citizen became customer allowed to cooperate or participate with initiatives. The social media benefits are (I) a better diffusion of contents through a cycle of sharing on the own blog and (II) the possibility for an open dialogue between municipalities and people that could express their opinion. The financial aspect underlined by the author is the opportunity to save agencies' costs spreading out the information through web.

In other articles more, specific application of where Public Administration could improve program as U.S. education are developed. (Lavertu, 2015) analysed the primary and secondary education in the public and private school of USA. Citizens have the aim to demand information about the private benefits that schools provide (opportunity and safety) opposite to information that captures the public goals (e.g., citizenship and equal opportunity). But the problem is that only few people have the knowledge or the possibility to assess the validity of performance metrics, introducing significant information failures in the educative market information. Government and non-profit organization supported by State should take seriously the sustainability of promotive campaign in improving valid information regarding the organization of public programs. The expansion of performance measurement to communicate a more accurate portrait of public value could by reach by big data utilization, that help to exploit a deeper knowledge of what are the educational mechanism and consequences associated to private and public benefits.

Anshari et al. (2016) has conducted a research on one of the most common source of data: smartphones. Given all the benefits the big data implies, the main scope is to understand how public sector could deploy an idiosyncratic use of technologies fitting the people desires. The main use of big data is for forecasting events and optimize process to support management data related to almost all the economic activities. It creates knowledge sharing across sectors, speed up the policy initiatives, is trigger for innovation and stimulates new opportunities. When people share, they also generate a huge amount of contents with smartphones. A key success factor is to develop Open

Government Data platforms that share real-time information with public, ensuring quick decision that will accelerate the business activities growth.

Table 13: Big data in Public Administration sector

Study	Potential research areas	Definition	Purpose
(Manyika, 2011)	Big data in European pubblic administration	big data as increasing transparency and applying advanced analytics.	The public sector could increase productivity and achieving higher levels of efficiency and effectiveness.
(Chen, 2012)	Field of security informatics and public safety	Sentiment analysis and data mining as a new tool	Develop of different BI&A framework to ensure the public safety against criminality and terrorism
(Lavertu, 2015)	Public administration accountability and efficiency	Big data assistant in improving the management of public programs.	"big data is already increasing the role of experts in performance measurement".
(Guillamón M. D., 2016))	Monitoring the interaction between citizens and PA	Big data and social media are a channel to communicate with people	Through Facebook the government can increase the transparency and reach more political consent/approval.
(Anshari M., 2016)	Big data application in e-government with data from smartphone	Big data is a source of people thinking	Proposal of a model with potential solutions of the big data in the e-government.

5.4. Big data in Manufacturing

Manufacturing is often referred to industry in which companies with high complex supply chain produce, trade and offer services or product. Data here are generated from sales department to inbound of material and machine sensors. So, the resources available are enormous and the possible outcomes not all explored. In the next lines a collection of what are the main applications, services, programs and major trend are discussed **Table 14**.

The supply chain of manufacturing firm supports different business activities and production tasks. Engineering, design, product development and post services are managed but the information belonging to this field remain separated. However, the utilization of big data tools could improve the communication between these distinct department with a creation of a PLM (product life cycle management) that integrates all the data to enable effective collaboration (Manyika, 2011), or the realization of platform that receives internal and external inputs to create new products. In particular in sector as Mobile operator, aerospace where substantially there is an assembly of different components this new platform could increase the communication and the fit between supplier.

In the manufacturing sector the application of big data is widespread and is giving even more attention. In this literature is included an article from Bumblauskas et al. (2017) that has realized a "Smart Maintenance Decision Support Systems (SMDSS)" based on big data concept. His study provides a predictive analytics model of supply chain breakdown that can be extended in the areas of quality, reliability and maintenance of manufacturer. This system is able to suggest the real-time status of the assets and the preventive measures that have to be performed on small power components, vehicles, miscellaneous, circuits etc. The difference between traditional and new big data predictive system of maintenance is in the analysis object. The traditional method is focused on monitoring time and condition, instead the predictive model simulates expected damage and failure. Some AS-IS application implies the use of apposite people, the maintainer, that collect data and supervises the machinery. With big data technology micro sensors change the structure introducing the remote collection of

information through WIFI connections. This sensor send input to an internal processor that monitor on time the status and elaborate its analysis.

The manufacturing system (MS) can be described with 5M:" Materials (their properties and functionalities), Machines (their task and precision), Methods (their productivity and efficiency), Measurement (their sensing) and Modelling (the ability to predict and optimize the process)" Jay Lee et al. (2015). Lee underlined how the advent of RFID, smart technologies allow a better collection of data coming from machinery but the doubt if these data provide the right information to the right location in the right time is still unexplored. In the article is defined the manufacturing information system (MIS) with 5C: Connection (between sensors), Cloud (physical place to store data), Content (meaning of information), Community (share the data) and Customization (personalize the value). MS and MIS are interrelated and coexist in the same organization. The problem referred to these two systems are themselves two: visible and invisible. Machine failure, time delays, product defect are some examples of visible issues and machine degradation, waste energy are the invisible one. To solve the previously cited problems a cyber-physical (coupled) model is realize, a real twin machine that operates in cloud with a simulation of conditions, using analytical algorithms as well as physical knowledge. This application is able to Sense, Store, Synchronize, Synthesis and Serve the right information helping the operational improvement of the manufacturing supply chain.

Opresnik et al. (2015) develop a theoretical framework to realize a connection between big data and "Servitization". This term is referred to the service-oriented strategy used by manufacturing companies in mature industries. It is an old practice to gain more competitive advantage and cut operation costs. The AS- IS servitization differs from the old one because of the technology effort and value in the delivery of Product-Services. Nowadays Tv are set with pay-per-view functionalities, ICT solutions supports business of everyday. But the point is that manufacturer are not owner of all the expertise and capabilities for sustain a high level of post services, therefore they collaborate each other combining the resources in a sort of Manufacturing Service Ecosystem. The article has the purpose to realize an integration between technology and initiatives to help the industries to increase the level of competitiveness. The evidence is that more manufacturer servitize its products or

goods and could obtain more information, and more data can exploit through the sharing to a third part.

Wang et al. (2016) develop a Maturity model for supply chain and logistic in the manufacturing sector. A maturity model is a framework that measures the ability of an organization to improve the quality of services in a specific area. In the literature review proposed by the author are analysed the relations between Business Analytics Big Data and operational performances. Five are the main capabilities on which build the model: functional, process-based, collaborative, agile Supply Chain Analytics and sustainable Supply Chain Analytics. It is related to the effective use of big data tool to achieve different goals. The supply chain has the problem to integrate the activities, to the duplication of processes and coordination between partners. To solve this first issue a functional SCA is used to analyses and solve problem belonging from the same activity. A process-based SCA help companies to solve process issue focusing on the distinctive steps and creating a common fit of operation. Collaborative SCA help the identification of key issues and activity sharing the relevant information, from internal and external data. Agile SCA allow companies to realign their resources according to the rapid changing in the external environment ensuring the competitivity. And the last solution is a Sustainable SCA that is defined as the use of big data to provide an appropriate information that can support decision making of sustainability issues.

Table 14: Big data in Manufacturing sector

Study	Potential research areas	Definition	Purpose
(Manyika, 2011)	All the business unit of the supply chain	BD as a source of operational value creation	Improve efficiency in design and production, promotion, distribution and quality.
(Bumblauskas D., 2017)	Supply chain activities of maintenance and	Smart Maintenance Decision Support Systems (SMDSS) based on corporate big data analytics	Creation of a system capable to improve asset lifecycle.
(Lee J., 2013)	Integration of IoT and production units of SC	Big data is a predictive tool to ensure SC performance	Framework creation that could help the reduction of visible and invisible issue of SC.
(Opresnik D., 2015)	Service oriented strategies	Big data as a strategy to support operation	Understanding the capability of BD to improve profitability in servitazion sector
(Wang G., 2016)	Logistic and supply chain control	BD as a tool to monitor performances	Review of tool and techniques to support the analysis of SC KPI

5.5. Big data in Retail and E-commerce industry

In retail sector there is a vastity of field in which big data is already introduced and where play a strategic role. Segmenting customer, forecast trends and managing supply chain are only few of the possible applications. Nevertheless, there are a lot of advantageous potential that could increase the competitive advantage of IT companies, particularly given the increase of the technological capabilities to store and analyse data. In addition, several subjects as suppliers, consumers, third parts are interfaced generating different information that could be processed to support decision making. For instance, five articles and real cases are described in order to summarize all the relevant initiatives adopted across this industry **Table 15**.

Big data have reshaped the activities in retail and e-commerce industries. Price comparation services and web-based markets have gain competitive advantage to some firms. The core business is no more physics retail as shops but are the web sites. As regard the firs innovation several companies provide a real-time pricing, creating major transparency for the final customers. Now it is possible to compare different vendors, saving costs with an average of ten percent. In addition, from retail perspectives, there is a cutting price to operate in this competitive market, an indirect attention on the operational efficiency in order to support this decreasing and the possibility to reinforce the visibility in sector (Manyika, 2011).

Kerstin Schäfer et al. (2013) conducts a research on the performance determination of web-sites and marketing analysis. As a result of internet, the presence on web becomes even more important, the mainly channel for communication that firm possesses. To address the importance of the online activities in the marketing sphere several approaches have been introduced, from web mining, click stream analysis and cookies monitoring. All of these techniques can be empowered by the application of big data tools changing the way in which the firm operate. One of the most important business activity is the application of relationship marketing strategies that are useful to develop, establish and maintain sustainable relation with the consumer and internet is becoming the main way to support these programs. Schäfer realizes a model in which are augmented the websites function with the related relationship management strategies. The first function is a commerce orientation to get the consumer that visit

the web site, in doing this a brand reputation is necessary to ensure repeated purchase. Then, the online platforms have the content and media provision function for the maintenance of the relationship and the support services to increase the customer satisfaction.

As it is already mentioned the use of internet continues to grow all around the world and the digital is becoming crucial for competitive advantage in both Business to consumer (B2C) and Business to business (B2B) marketing. Leeflang et al. (2014) introduce the new market Consumer to consumer (C2C) that with the use of web is emerging, example of companies operating in this sector are eBay, Lulu etc. As the internet giving more attention creating more value, some risks relating to his use are growing. In this field the reputation is crucial, the study suggests that 90% of consumers read online reviews before to choose the product and 70% read also the first four lines of comments. The studies identify four main challenges in the marketing area that could be managed with big data support. First to compete effectively is important to understand the customer insights and analyses their data. Second, understand and accept the power of social media for brand equity and customer relation. Third, analyse what are the new digital metrics and asses the new digital marketing activities, and in the end increase talent gap in analytic capabilities.

Lau et al. (2016) defines big data commerce as a data centric markets, an extension of traditional e-commerce that embraces new theories, systems, models and methods that help companies in the amount of data generated from different sources. In the definition of big data chapter, the 5V are discussed and supported by real cases. In this article the author connects the 5Vs to five typical arguments of ecommerce sector, the 5P: product, price, people, promotion and place. With the analysis of clickstreams and texts is possible develop dynamic pricing strategies in real time. Business activities already used by airplane companies, or hotel retailer etc. to address more value to online products. Products represent the second P, because monitoring customer transaction and satisfaction is possible to perform even more addressed goods or services, more customized that better fit with consumer preferences. This is made possible through the implementation of new machine with algorithm able to understand consumer lifestyle. With sensors, mobile devices and wearable technologies is becoming easier to design sophisticated representation of new

interesting markets. According to the age, the behaviour, the health conditions and current emotions will be possible also to give to the (specific) customer the right promotion to the right time and place. The last P is people, that is the word used to describes the customer profile generated from these new technologies. All the previous actions and understanding cited are interconnected and filter to realize a sort of identikit that is consulted when programs or new initiatives will be activated.

As it is already discussed, the big data peculiarity is to predict or support decision making action. Matthias et al. (2017) presents two real case studies regarding retailer company that have experienced with the use of primary data, that are generated internally to the firm. The first case has as protagonist an online white goods retailer, with the aim to tailor the experiences for the consumer that visit web site. In internet, the lack of physical aspects as age, gender, and in general the appearance makes difficult a categorization of the consumer. So, the raised question is: could the historical click actions give more addressed information about the consumer? The research is conducted comparing different variables as the smartphone use to surf, the OS (operation system) and the browser. The evidence is referring to a strong association between operation system and browser combination. People with safari or google are typical owner of devices with that default browser, and generally express preferences for the same products. As result the big data identify two customers category based on their use of browser: (i) more interested in traditional solutions, (ii) more interested in modern. The second case study tells about a multi-channel UK retailer that share his own business in three macro categories: inbound, re-sales and retail. During the activities years it has collects a huge amount of data and is ready for the analysis using scientific platform. The process shows a strange behaviour of goods sold with guarantee, that appear with more appeal but less profitable. What the cases learn is that Big data have the promise to gain information hidden on the own data. In particular the case 1 highlights the chance to understand consumers' preferences to perform right product, and case 2 could be used to describe how Big data could redesign the portfolio of services and goods according to the profitability and the differentiation.

Table 15: Big data in retail sector

Study	Potential research areas	Definition	Purpose
(Manyika, 2011)	Implications on marketing, operational performances and information flow in US retail	BD as a source of business value creation	Improve transparency, segmentation and replacing/supporting human decision.
(Schäfer K., 2013)	Information technology architecture and marketing functionalities	An evaluation instrument to support business decision	A framework to analyse website activities and develop different approach
(Leeflang P. S.H., 2014)	Functional division: Supply Chain, organizational structure and brand management.	BD as a tool to improve performances	Reduce "talent gap", redesign organization and develop new metrics.
(Lau R Y.K., 2016)	Focus on strategic and market decision in e-commerce environment.	Big data is a tangible capital with the traditional financial and intellectual capitals for businesses	Addressing 5V of BD to 5P of BD commerce.
(Matthias O., 2017)	Retail supply chain	BD is view as an improvement factor for operational performance	The objective is to understand if organisations use BD and if its benefits match its expectative.

5.6. Framework application among industries

The research points out a classification of vertical papers that have analysed the dimension of the general framework **Figure 13** and **Figure 14**. In particular two matrices are developed to connect (I) tangible, intangible and human resources with application performed in the five industries; (II) five dimensions strictly associated to the firm performances with the sector structure and objective. The dimensions are collected and individualized according to the most important benefits gathered from firm through the application of big data **Table 16** and **Table 17**. They are: Creating transparency, operation efficiency, segmenting populations to customize actions, supporting human decision making with automated algorithms, innovating new business, models, products, and services. As regard the financial performance this technology is able to ensure more transparency between the two parts involved in the interactions, resulting in more trust and improvement of the exchange. In this intersection the most involved sector is the financial one. The operational performance is related to the operational improvement of the efficiency with chances to save cost, better perform product, or services, in the right place and time. *And* also supporting human decisions with algorithm is viewed as a dimension of operation efficiency. The manufacturing industries are particularly doing more effort in efficiency of their assets. The fourth dimensions individualized in the table is segmenting populations to customize actions and it is connected with strategic performance, retail and manufacturing. They are particularly interested in reach this capability to ensure a better position and competitive advantage. The last dimension is innovating new business models, products, and services that relates to the ability to forecast with the founded pattern in the dataset. Also, in this sector companies like Procter and gamble, Amazon, Alibaba are doing the major effort and are gaining even more advantage in the market. The matrix shows also connection not analysed or developed. In particular the associations of public administration with supporting human decision making with automated algorithms, healthcare and manufacturing with segmenting populations to customize actions have lack of articles or initiatives. This is not evidence of not relationship between the fields but has to be input for further research suggestions and exploration

Table 16: Framework Performances across industries

		Sectors					
Dimensions	Firm performance	Healthcare	Financial	Public administration	Manufacturing	Retail e-commerce	
Creating transparency	Financial performance	Manyika et al. (2011);	Gui et al. (2015), Jiang et al. (2014); E. Clemons (2017);	Manyika et al. (2011), Guillamón et al. (2016)	Jay Lee et al. (2015); Wang et al. (2016)	Manyika et al. (2011);	9
Operation efficiency	Operational performance	Jing Wu (2016);	Turner et al. (2013);	Chen et al. (2012); G Kim (2014);	Manyika et al. (2011); Jay Lee et al. (2015); Bumblauskas et al. (2017); G.J. Hahn (2015)	Lau et al. (2016); Schäfer et al. (2013)	10
Supporting human decision making with automated algorithms	Operational performance	Pramanik (2017); Wang et al. (2016)	Chan et al. (2016); Kim (2016),		G. Wang et al. (2016); Bumblauskas et al. (2017); G.J. Hahn (2015)	Manyika et al. (2011); Lau et al. (2016); Leeflang et al. (2014)	10
Segmenting populations to customize actions	Strategic performance		Gui et al. (2015); E. Clemons (2017);			Matthias et al. (2017); Lau et al. (2016)	4
Innovating new business models, products, and services	Marketing performance	Pramanik (2017); Wang et al. (2016); Chen et al. (2012); Jing Wu (2016);	Seddon et al. (2016);	Anshari et al. (2016); E. Clemons (2017);	Oprensik et al. (2015); Zhan (2017),	Matthias et al. (2017); Pousttchi et al. (2014); Zhan et al. (2017), E. Clemons (2017);	12

As regard the intersection between resources and sectors, **Table 19** shows how almost all the antecedents of big data capabilities are necessary to gain the improvement categorized in the previous table. The curiosity of the matrix is in the number of articles across sector that use technical skills. In the framework it is well mentioned the importance of having skilled personnel like data analyst, manager, data scientist to success in the big data application but here is the most important antecedent. All the application and program developed by financial, public administration, healthcare, manufacturing and retail companies have as prerequisites employs well prepared in big data capabilities. The second most utilized resource is the technology such as

Hadoop, clouds and other programs that are useful first to support human decision with data and then to prevent disease or breakdown especially in the manufacturer sector.
As regard the aspect not investigate from literature in **Table 17** and **Table 18** are collected and summarize the main findings. The question is, all the aspects not founded, are not necessary for the specific sector or simply there are no studies with the aim to investigate the topic? Trying to answer these questions, some researches argue that Healthcare and Public administration rarely segmenting population to customize action because they ignore the potential of big data in improving the public welfare. But also, are less interested in supporting human decision with automated algorithms preferring the guide of politicians knowledge.

Table 17: Framework performances not investigated among sectors

Sectors	Framework Variables	Lack of references	Reason
Healthcare	*Strategic performance*	Segmenting population to customize actions	Healthcare ignores the potential of segmenting population applied to public welfare
Public Administration	*Operational performance*	Supporting human decision making with automated algorithms	PA business structure is guided by politicians knowledge and decisions
	Strategic performance	Segmenting populations to customize actions	PA ignores the potential of segmenting population applied to public welfare

As regard the resources, financial sector and retail have lack of intangible, tangible and human assets, specifically they retain not necessary basic resources, managerial skills and organizational learning. This lack of references probably is due to two main reasons, first of both sector still be in the "introduction" phase of big data, so have no necessity to sustain programs and initiatives through these human and intangible resources. Second, basic resources as time and investment are included in other resources, hire capable human require investment as teach the work to people require time.

Table 18: Framework resources not investigated among sectors

Sectors	Framework Variables	Lack of references	Reasons
Financial	*Tangible resource*	Basic resources	Not considered as independent resources because included in other resources
	Human resource	Managerial skills	Financial sector is still in the «introduction» phase of Big Data
	Intangible resource	Organizational learning	Financial sector is still in the «introduction» phase of Big Data
Retail e-commerce	*Tangible resource*	Basic resources	Not considered as independent resources because included in other resources
	Human resource	Managerial skills	Retail e-commerce is still in the «introduction» phase of Big Data
	Intangible resource	Organizational learning	Retail e-commerce is still in the «introduction» phase of Big Data

Table 19: Framework Resources across industries

		Sectors				
Resources	**Typology**	Healthcare	Financial	Public administration	Manufacturing	Retail e-commerce
Data	Tangible	(Manyika, 2011; Wu J., 2016)	(Turner D., 2013)	(Manyika, 2011) (Anshari M., 2016)	(Bumblauskas D., 2017)	(Matthias O., 2017)
Technology	Tangible	(Manyika, 2011; Chen, 2012)	(Gui & Ming, 2015; Chan S. W.K., 2016; Seddon J.J.M., 2017)	(Anshari M., 2016)	(Lee J., 2013; Bumblauskas D., 2017; Hahn G.J, 2015)	(Lau R Y.K., 2016)
Basic resource	Tangible	(Pramanik Ileas, 2017)		(Anshari M., 2016)		
Technical skill	human	(Chen, 2012; Pramanik Ileas, 2017; Wang, 2016)	(Chan S. W.K., 2016; Turner D., 2013; Seddon J.J.M., 2017)	(Chen, 2012)	(Hahn G.J, 2015; Bumblauskas D., 2017)	(Leeflang P. S.H., 2014; Schäfer K., 2013)
Managerial skill	human	(Chen, 2012; Wang, 2016)		(Kim G., 2014)	(Opresnik D., 2015)	
Organizational learning	Intangible	(Wang, 2016)		(Kim G., 2014; Janssen M., 2017)	(Manyika, 2011; Wang, 2016)	
Data driven culture	Intangible	(Wu J., 2016)	(Turner D., 2013)	(Manyika, 2011; Janssen M., 2017)	(Bumblauskas D., 2017; Hahn G.J, 2015)	(Lau, 2012; Matthias O., 2017)

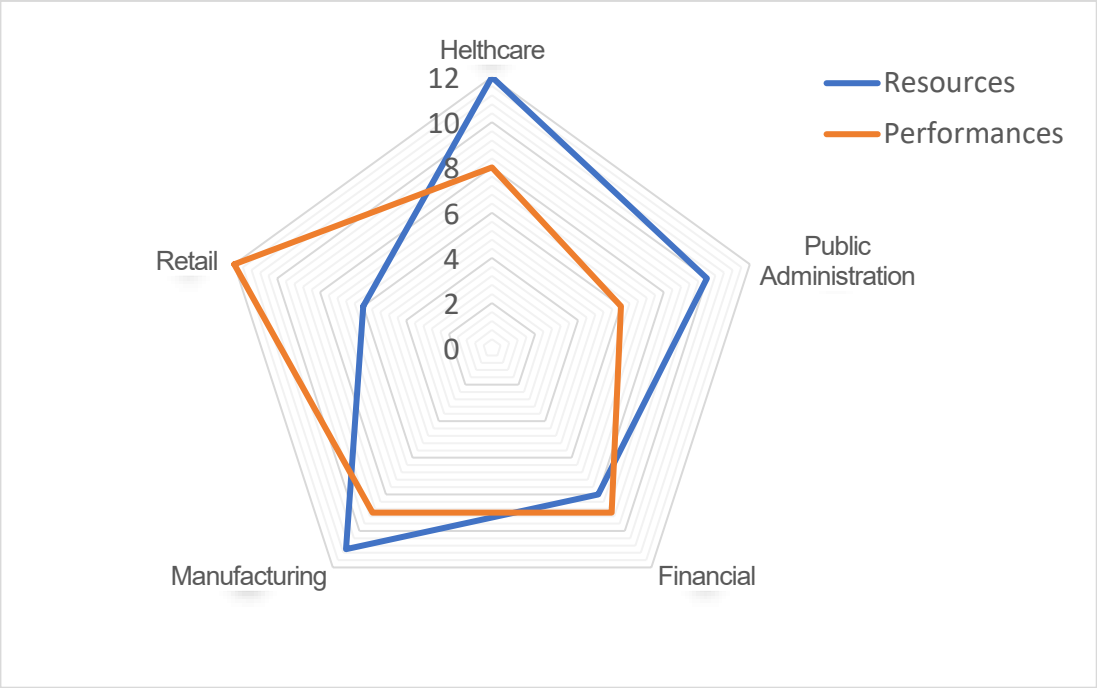

Figure 13: # of references with framework application

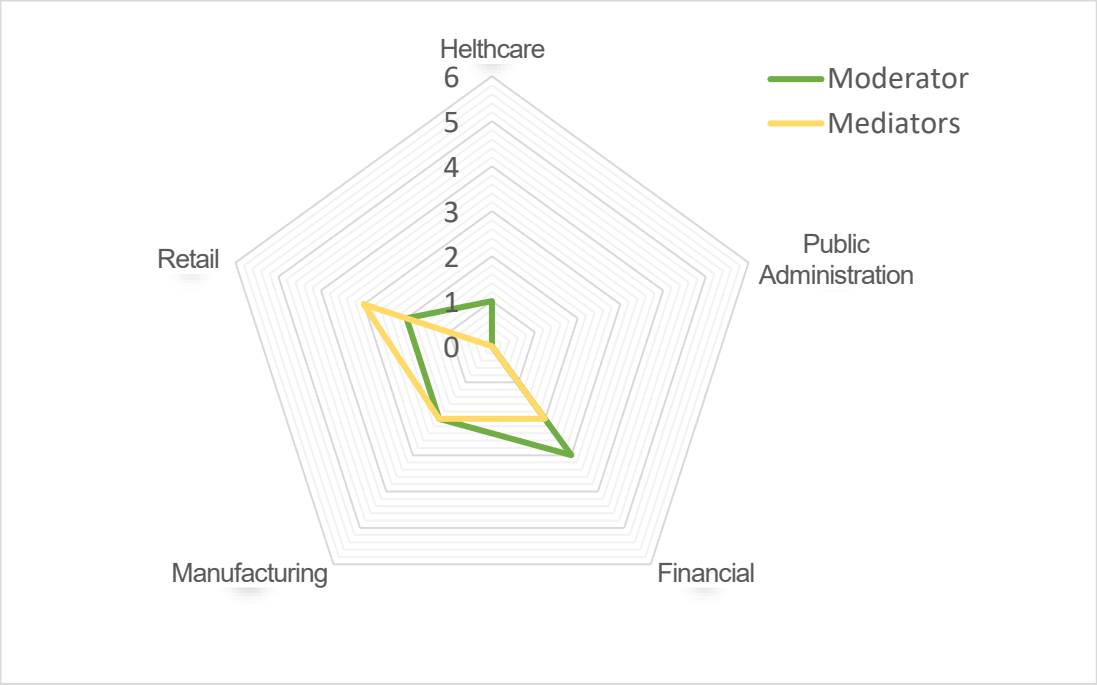

Figure 14: # of references with integrated framework application

6. Research agenda

In this chapter future directions are suggested according to the theme developed in the previous paragraphs. Firstly, some unexplored ways on the framework realization are described, in order to validate and cover more issues. And then given the novelty of the topic, future extension of industries analysis is indicated to ensure a deeper understanding of the state of art.

6.1. Suggestion for next frameworks

The chapter five presents a research model of the big data impact on firm's performances, built on the capabilities of different companies in the managing this new technology. The principle of aggregation and framework construction is developed tracking the relation between big data capabilities and performances translated in operational, financial, market and strategic efficiency.

It is performed all the value chain of big data, generalized by De Mauro A., 2016 in four main steps: information, technology, methods and impact. Information is referred to knowledge that comes from "datafication" (Mayer-Schönberger V., 2013), the process that converts into a digital format to be organized into more structured data sets. Technology and methods are referred to the two capabilities, human skills and real assets. As regards the methods, it is clear that several sectors with several problems need to be solved. It is previous shown that big data have a consistent level of adaptability for each case.

Firms spend millions of dollars on business analytics to enhance business value and performance. However, studies on business analytics to business outcomes show mixed results (Ren S. Ji-fan, 2016). In this study big data are shown as inhibitor of positive effect on firm performances, ensuring business alignment and dynamic capabilities. Big data seen as engine, tool of value creation and source of competitive advantage (Côrte-Real N., 2017). It is highlighted positive growth of the economic indexes (Akter S., 2016), of the risk mitigation, of the supply chain visibility, of the asset productivity (Gunasekaran A., 2017). However, the literature covers only a small

part of the context and field of application. The landscape expected involves researches and studies that are trying to connect firm resources and performances underlining the effective correlations. The 'big data' academic studies have a very strong focus on opportunities. Far less attention has been paid to the threats that arise from data. There Are possible negative outcomes, already introduced by (Clarke, 2015) ,that need more attention and development to be prepared for unexpected events. All the studies analysed in chapter five underline positive effects between independent and dependent variables, with the automatic deduction that is always a positive choice the shift in data drive decision making (McAfee & Brynjolfsson, 2012). Is it always true or sometimes is it better focusing on own resources without following the innovations? And what is the effective investments that firms must allocate to overreach the real benefits? Summarizing, suggestion for future research and studies can be grouped in three initiatives: (i) what are the negative effects of the adoption of big data on firm performances, what is the dark side in managing this technology; (ii) what are the relations not individualized and that must be considered, both resources and performances side and (iii) if exist more or different mediator and moderators between big data capabilities and firm performances.

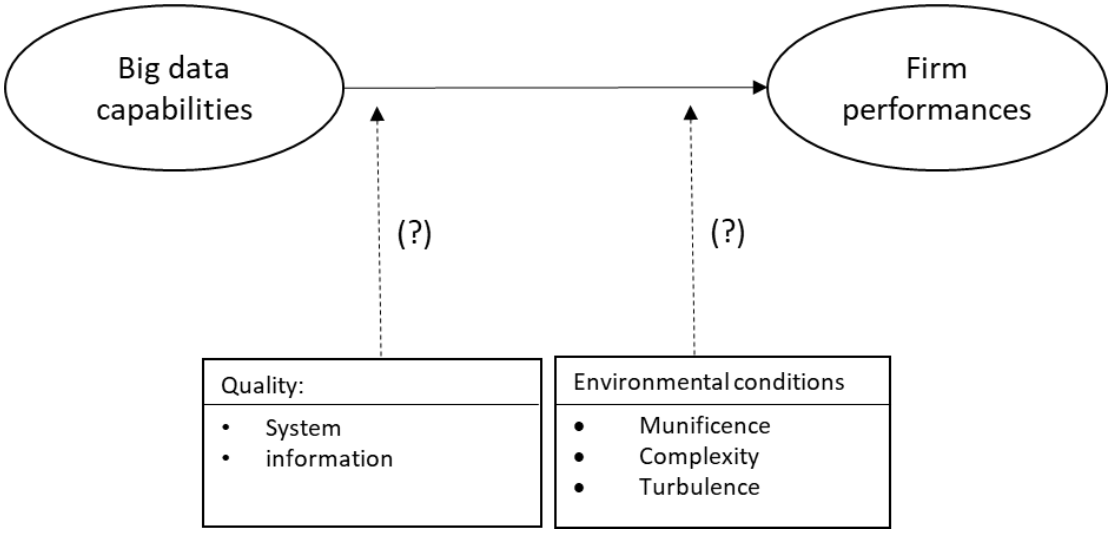

Figure 15: Future suggestion of moderators

As regard the last point, several studies of Information Technology capabilities have discovered other variables: environmental and quality (**Figure 15**). Environmental condition implies 3 different scenarios, environmental dynamism/turbulence, environmental munificence and environmental complexity. Both of these aspects moderate negatively the impact of the it capabilities on FPER due the realignment of technological configuration (Neirotti P., 2017; Wade M., 2004; Schilke O., 2014). But other position consider Big Data Analytics use as salient in highly dynamic environments because Big data capabilities is expected to increase an organization's capacity to discover new knowledge and insight (Chen D. Q., 2015).

And the quality of the information, of the system could improve the impact on FPER or it is a prerequisite of tangible resources? (Ren S. Ji-fan, 2016; Wieder B., 2015).

6.2. Future direction

Big data enables firms' operation improvements at many levels. While the current study analysed some of them, there are still many unanswered questions that require further works (Wamba S. F. F. R., 2017). The objective of this chapter is to give support to the successful big data analysis, to provide direction for future researches especially in the industries development.

This study examines the relation of big data analytics across sectors, categorizing their effects and consequences. Chapter 6 can be implemented adding and discovering new solutions adopted by companies, understanding what the AS-IS solutions are and what the possible TO-BE implementation. How the firms are and how they are transforming their businesses and assets to reach competitive advantage but also in organizations where is important the social benefits as public administration and healthcare. Business intelligence is seen as a new way to reduce the gap among firms in manufacturing sector (Manyika, 2011). Clustering, customizing products and addressing new services (Chen, 2012) are the current challenges for the retail industries. Healthcare and public administration are upgrading their system ensuring the alignment within this context raising the service level (Lavertu, 2015). For example, how organizations can use this data to improve services or add data-enabled enhancements to their products needs to be investigated (Wamba S. F. F. R., 2017).

Studies that encompass all cross functional facets for better understanding and growth of knowledge in this domain are essential (Wamba S. F. S. A., 2015). New insights in relation to management are also necessary: how can the firm adopt this technology supporting their core business? The five challenges of McAfee could be integrated and discussed, companies won't reap the full benefits of a transition to big data unless they're able to manage change effectively (McAfee & Brynjolfsson, 2012).

Technologies is always upgrading and in evolution, creating positive effect of renew and negative effects for firms that need to follow the change to ensure positive value of the investments. (Katal A., 2013). Issues that have to be analysed in detail, to explore the effective innovation Pareto's line. It is interesting to find that not much focus is present in existing literature to discuss the platforms and tools which enable the outcome realization from big data problems (Grover P., 2017).

Big data still aims in large part to deliver the right information to the right person at the right time in the right form but is now able to do so in a significantly more sophisticated form that required more attention and classification (Agarwal S., 2014). Summarizing, the challenges and topics in which concentrate the future efforts are (i) collect more real cases study to better perform the structure of each sector in order to understand the prerequisites needed for the big data implementation; (ii) extend the analysis on the sector not included in this research, how the data-driven companies, such as Uber, Airbnb, have been completely reshaping the ways of doing traditional business in automotive, transportation, retail, and consumer goods industries; and (iii) applied the framework in chapter five on all the industries trying to realize also a quantitative report to test empirically the relation. Further analysis can only help the creation of an accurate literature body to ensure a deeply understanding of this articulated topic.

7. Conclusions

The research provides a collection of 212 articles, that could be a starter point to explore more in detail all the issues presented. It gives different suggestion to develop open points and investigate the topic with both an economic and scientific perspective. The sphere of researching is in management, economics and information field. Obviously big data is directly connected to informatic issues, regarding for example algorithms, program languages and architecture management. Strategical and organizational implications are only at the beginning of the research or of the application. For these reasons, the aim of this work is to support and further investigate the management consequences in dealing with this technology.

The research is articulated in three main sections. The first one is referred to the introduction of big data concept and to the description of their role in the society. The huge amount of information daily generated at high rate are gaining attention both from economic and scientific studies.

The second section is the core of the research. It is a collection and elaboration of thirty articles, that present frameworks of big data model with the relations between dependent and independent variables. In particular, the framework in **Figure 10** individualizes resources, big data capabilities and firm performances connecting them with a rich network of positive links. The resources are classified according to the Resource Based View in tangible, intangible and human. All the seven resources (data, technology, data driven culture, basic resources, organizational learning, human skill, technological skill) are antecedents of three big data components: big data technological capability, big data management capability, big data human capability. Together the three capabilities have positive impact on firm performances. Specifically, for this research the KPI (Key Performances Indicator) considered are belonging to financial, operational, strategical and marketing field. The study was than integrated with moderators and mediators variables as agility, business alignment, marketing capability, analytics capability and top management commitment. The introduction of these variables does not change the positive impact previously recognized, but in some cases, they empower it.

The last section is dedicated to the application of the research model across the industries. The effects of big data capabilities are analysed on five main sectors: healthcare and wellbeing, financial, public administration, manufacturing and retail e-commerce, that include all the possible big data applications. The analysis shows that technical skill (Human Resources) and technology (Tangible Resources) are the most common resources utilized across both five sectors and that analytics (Human Resources) and managerial skill (Intangible Resources) are useful for all the industries. Innovating new business models, products and services is the most cited dimension due to the forecast capabilities of big data necessary for all the sectors. This trend is preeminent in retail companies that are gaining competitive advantage thanks to this technology.

In conclusion it is important to underline that this Big Data capabilities research model should not be considered as a universal framework. The literature is only at the beginning of understanding the big data phenomenon, and therefore building a satisfying collection of organizational resources connected to the creation of Big Data capability is not easy.

8. References

Abadi D., R. A. (2016). The Beckman Report on Database Research. *Comunication of ACM*, 92-99.

Abbasi A., S. S. (2016). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*.

Abbasi A., S. S. (2016). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*.

Addo-Tenkorang R., P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering*, 528-543.

Agarwal S., V. D. (2014). Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Information Systems Research*, 443-448.

Agarwal, R. (2012). The benefits of combining data with empathy. *MIT Sloan Managemet* , 35-41.

Akter S., S. F. (2016). How to improve firm performance using big data analytics capability. *Int. J. Production Economics*, 113-131.

Andrea De Mauro, M. G. (2016). A formal definition of Big Data based. *Library Review*, 122-135.

Anshari M., S. A. (2016). E-Government with Big Data Enabled through Smartphone for Public Services: Possibilities and Challenges. . *International Journal of Public Administration*, 1143-1158.

Barney J. B. & Hesterly S. W. (2011). *Strategic Management and Competitive Advantage.* Prentice Hall PTR.

Boyd D., K. C. (2012). CRITICAL QUESTIONS FOR BIG DATA. *Information, Communication & Society*, 662-679.

Bughin, J. (2016). Big data, Big bang? *Journal of Big Data*, 1-14.

Bulent Menguc, S. A. (2006). Creating a Firm-Level Dynamic Capability through Capitalizing on Market Orientation and Innovativeness. *Journal of the Academy of Marketing Science*, 63-73.

Bumblauskas D., D. G. (2017). Smart Maintenance Decision Support Systems (SMDSS) based on corporate big data analytics. *Expert Systems With Applications*, 303-317.

Chan S. W.K., M. W. (2016). Sentiment Analysis in Financial Texts. *Decision Support Systems*.

Chasalow, L. (2015). Factors Contributing to Business Intelligence Success: The Impact of Dynamic Capabilities. *Twenty-first Americas Conference on Information Systems*, (pp. 1-14). Puerto Rico.

Chen D. Q., D. S. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems,*, 4-39.

Chen, H. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 1165-1188.

Clarke, R. (2015). Big data, big risks. *Information System of Journal*, 77-90.

Constantiou I.D., J. K. (2014). New games, new rules: big data and the changing context of strategy. *Journal of Information Technology*, 1-14.

Côrte-Real N., T. O. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 379-390.

Cosic R., S. G. (2012). Towards a Business Analytics Capability Maturity Model. *23rd Australasian Conference on Information Systems.*

Cukier K. (2010). Data, data everywhere: A special report on managing information. *The Economist*.

Davenport T. (2014). *Big data at work: Dispelling the myths, uncovering the opportunities.* Harvard Business Review Press.

Davenport, T. (2012). The Human Side of Big Data and High-Performance Analytics. *Internationa lInstitute for Analytics*, 1-13.

Day j., w. l. (2011). Catching the long tail: competitive advantage through distribution strategy. *Journal of Hospitality advantage through distribution strategy*, 204-214.

Dijcks, J.-P. (2013). *Big Data for the Enterprise.* oracle.

Dubey R., G. R. (2017). Can big data and predictive analytics improve social and environmental sustainability? *Technological Forecasting and Social Change*.

Eckstein, D. G. (2015). The performance impact of supply chain agility and supply chain adaptability: The moderating effect of product complexity. *International Journal of Production Research*, 3028-3046.

Erevelles S., N. F. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 897-904.

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

Gang-Hoon Kim, S. a.-H. (2014). Big-Data Application in the Government sector. *Comunication of ACM*, 1-8.

Grant R. M. (1991). *Contemporary Strategy Analysis.* John Wiley & Son, Ltd.

Grover P., A. K. (2017). Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature. *Global Journal of Flexible Systems Management*, 203-229.

Gui, F. Y., & Ming, Z. J. (2015). Analysis on Information Value of Big Data in Internet. *International Conference on Logistics, Informatics and Service Sciences (LISS)*, (pp. 1-6).

Guillamón M. D., A.-M. R. (2016). Factors influencing social media use in local governments: The case of Italy and Spain. *Government Information Quarterly*, 460-471.

Gunasekaran A., T. P. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*. doi:http://dx.doi.org/10.1016/j.jbusres.2016.08.004

Guoqing Chen, X. G. (2016). Big Data Commerce. *Information and Management*.

Gupta M., J. F. (2016). Toward the Development of a Big Data Analytics Capability. *Information and Management*.

Hahn G.J, J. P. (2015). A perspective on applications of in-memory analytics in supply chain management. *Decision Support Systems*, 45-52.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance.

Halevi, G. M. (2012). *The Evolution of Big Data as a Research and Scientific Topic.* Retrieved from www.researchtrends.com: http://www.researchtrends.com

Hassna G., P. B. (2016). Big data capability, customer agility, and organization performance: A dynamic capability perspective. *Theory Development Workshop, International Conference on Information Systems.*

Helfat C.E. & Peteraf M.A. (2003). The dynamic resource-based view: capability lifecycles. *Strategic Managemant Journal*, 997-1010.

Jagadish H.V., J. G. (2014). Big Data and Its Technical Challenges. *COMMUNICATIONS OF THE ACM*, 86-94.

Janssen M., H. v. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 338-345.

Jie Sheng, J. A.-A. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, 97-112.

Jonathan J.J.M. Seddon, W. L. (2016). A model for unpacking big data analytics in high-frequency trading. *Journal of business Research*, 300-307.

Kache F., S. 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*, 10-36.

Katal A., M. W. (2013). Big Data: Issues, Challenges, Tools and Good Practices. *Contemporary Computing (IC3.* Noida, India: IEEE.

Khan, S. &. (2017). A survey on scholarly data: From big data perspective. *Information Processing & Management*, 923-944.

Kim G., S. a.-H. (2014). Big-Data Applications in the Government sector. *communications of the acm*.

King, R. (2012). Ford gets smarter about marketing and design. *Wall Street Journal*.

Kiron, D. P. (2014). The analytics mandate. *MIT Sloan Manag*, 1-25.

Krugler, L. (2016). What Happens When Big Data Blunders? *Comunication of ACM*, 15-16.

Kung L., A. J.-F. (2015). Managing Big Data for Firm Performance: a Configurational Approach. *Twenty-first Americas Conference on Information Systems.* Puerto Rico.

Kwona O., L. N. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 387-394.

Lau R Y.K., J. L. (2016). Big Data Commerce. *Information and Management*.

Lau, R. L. (2012). Web 2.0 environmental scanning and adaptive decision support for business mergers and acquisitions. *MIS quarterly*, 1239-1268.

Lavertu, S. (2015). *We All Need Help: "Big Data" and the Mismeasure of Public Administration.* Public Administration review.

Lavertu, S. (2015). We All Need Help: "Big Data" and the Mismeasure of Public Administration. *The American Society for Public Administration*.

Lee J., E. L.-a. (2013). Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters*, 38-41.

Leeflang P. S.H., P. C. (2014). Challenges and solutions for marketing in a digital era. *European Management Journal*, 1-12.

Manyika, J. C. (2011). *Big data: The next frontier for innovation, competition and producivity.* Mckinsey Global Institute.

Marr, B. (2015, september 30). *Big Data: 20 Mind-Boggling Facts Everyone Must Read.* Retrieved from www.forbes.com.

Matthias O., I. F. (2017). Making sense of Big Data: can it transform operations management? . *International Journal of Operations & Production Management*, 37-55.

Mayer-Schönberger V., C. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think,.* London.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*.

Meyer-Waarden Lars, S. S. (2014). *Big data resources, marketing capabilities, and firm performance: the moderating effect of choice ogf business strategy.* European Union's Horizon 2020 research and innovation programm.

Neirotti P., R. E. (2017). On the contingent value of IT-based capabilities for the competitive advantage of SMEs: Mechanisms and empirical evidence. *Information & Management*, 139-153.

Nguyen T., L. Z. (2017). Big data analytics in supply chain management: A state-of-the-art. *Computers and Operations Research*, 1-11.

Nitzl C., J. L. (2016). Mediation Analysis in Partial Least Squares Path Modeling: Helping Researchers Discuss More Sophisticated Models. *Industrial Management & Data Systems*.

Olszak, C. M. (2016). Toward better understanding and use of business intelligence in organizations. *Information Systems Management*, 105-123.

Opresnik D., M. T. (2015). The value of Big Data in servitization. *Int. J.ProductionEconomics*, 174-184.

Orlikowski W. J. (2007). Sociomaterial Practices: Exploring Technology at Work. *SAGE Journal*.

Peteraf M. A., B. (2003). Unraveling there source-based tangle. *Managerial Decisions Economics*, 309-323.

Peteraf, M. (2003). Unravelingtheresource-basedtangle. *Managerial and Decisions Economics*, 309-323.

Porter M. E. (1985). *Competitive Strategy.* New York: The Free Press.

Pramanik Ileas, R. Y. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems With Applications*, 370-383.

Ren S. Ji-fan, S. F. (2016). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*. doi:10.1080/00207543.2016.1154209

Rubin, V. L. (2012). Discerning truth from deception: Human judgments and automation effort. *First Monday*. Retrieved from first monday.

Schäfer K., T.-F. K. (2013). Determining the performance of website-based relationship marketing. *Expert Systems with Applications*, 7571–7578.

Schilke O. (2014). On the contingent value of dynamic capabilities for competitive advantage: The nonlinear moderating effect of environmental dynamism. *Strategic managemant journal*, 179-203.

Seddon J.J.M., W. L. (2017). A model for unpacking big data analytics in high-frequency trading. *Journal of Business Research*, 300-307.

Shanks, G. &. (2011). Creating value from business analytics systems: the impact of strategy. *15th Pacific Asia Conference on Information Systems: Quality Research in Pacific* (pp. 1-12). Queensland University of Technology.

Sharma R., S. M. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*.

Singh A., D. J. (2016). Application of Big Data in Supply Chain Management. *5th International Conference of Materials Processing and Characterization.*

Tambe P. (2014). Big data investments, skills and firm value. *Management Science*, 1-36.

TechAmerica Foundation's Federal Big Data Com. (2012). *Demystifying bigdata: A practical guide to transforming the business of Government.* Retrieved from http://www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf

Teece D. J. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 509-533.

Thibeault, J. W. (2014). *Delivering Digital Experiences That People Want to share.* John Wiley & Sons.

Turner D., M. S. (2013). *Analytics: The real-world use of big data in financial services.* Oxford: IBM Institute for Business Value.

Vidgen, R. ,. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, 626-639.

Wade M., H. J. (2004). The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. *MIS Quartely*, 107-142.

Wamba S. F., F. R. (2017). Transforming operations and production management using big data and business analytics: future research directions. *International Journal of Operations & Production Management*, 1-9.

Wamba S. F., S. A. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *Int. J. Production Economics*, 234-246.

Wamba S. Fosso, A. G.-f. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 1-10.

Wang G., A. G. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *Int. J. Production Economics*, 98-110.

Wang, Y. e. (2016). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting & Social Change*.

Wieder B., M.-L. O. (2015). The impact of Business Intelligence on the quality of decision making – a mediation model. *Conference on ENTERprise Information Systems.* Broadway: Procedia Computer Science.

Wixom B. H., H. W. (2007). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly*, 17-41.

Wu J., H. L. (2016). The Promising Future of Healthcare Services: When Big Data Analytics Meets Wearable Technology. *Information and Management*.

Yaqoob I., &. H. (2016). Big Data: From Beginning to Future. . *International Journal of Information Management.*

Zhao J.L., S. F. (2014). Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics*, 169-174.

Ziora, A. C. (2015). The role of big data solutions in the management of organizations. Review of selected practical examples. *International Conference on Communication, Management and Information Technology*, (pp. 1006-1012). Czestochowa, Poland.