UNIVERSIDAD DEL ROSARIO

MONTPELLIER BUSINESS SCHOOL

Impact of big data in organizations and

customer relationship

Bachelor Thesis

Carlos Daniel Velasco Yara

Bogotá

2020

UNIVERSIDAD DEL ROSARIO

MONTPELLIER BUSINESS SCHOOL

Impact of big data in organizations and

customer relationship

Bachelor Thesis

Carlos Daniel Velasco Yara

Phillipe Giuliani

Administración de empresas

Bogotá

2020

Tabla de contenido

Resumen

Esta investigación tiene como objetivo mostrar la incidencia de Big data cómo fenómeno dentro de nuestra sociedad en general, pero centrándose en su incidencia a nivel organizacional y en la relación con los clientes. Todo esto gracias al análisis de la literatura de importantes autores de diversos campos que junto con las entrevistas y recolección de datos ayudarán a comprender mejor los cambios masivos que este fenómeno está provocando, considerando a través de este texto sus potenciales riesgos y oportunidades que determinarán los verdaderos beneficios para las organizaciones y su posición dentro del mercado y la sociedad en días donde la adquisición de tecnología por si sola ya no es fuente de ventaja competitiva sino que su uso eficiente significará una posición privilegiada.

Palabras Clave

Big data, Stakeholder, ciencia de datos, Internet de las cosas, negocios digitales

Abstract

This research aims to show the incidence of Big data in the whole but focusing mainly in the organizations and customer relationship. All of this thanks to the analysis of various authors literature from diverse fields that in combination with interviews and questionnaires will help to look into better understand the massive changes that this phenomenon is sparking , considering their potential risks and opportunities to be regarded as a very useful tool for organizations in this days were the single technology acquisition is no longer a source of competitive advantage but instead their efficient uses will mean a privileged position .

Keywords

Big data, technology, stakeholder, organizations, knowledge, science, data science, internet of things, digital business

Impact of Big Data in organizations and customers relationship

1.   **Literature Review**

Big data benefits

(Grable & Lyons,2018) Stablish in their columns that "big data can provide more precise insights into hidden patterns, trends, and associations, especially in the context of human decision making." Setting a new way to treat customer data with more accurate analysis, but in their text: Introduction to Big data they mentioned that nowadays this detailed analysis became possible thanks to four statistical techniques and methods as Artificial neuronal network (ANN) "Whose strength lies in the creation of learning algorithms that find patterns that have predictive power", often with the involvement of historical trends and transactional data to predict future events and make predictions.

Data mining

Is defined by the authors as the use of already existing information and analysis of data to "search for hidden or emerging patterns in data to explain a particular phenomenon. This statistical technique is often used to classify sequences and describe data, allowing to describe detailed future trends thought clustering models (the method to set group objects to find similar characteristics of the individuals in the group), decision trees and neural networks.

This method has several applications in organizations, allowing to develop risk analysis trough pattern associations "For example, credit card issuers use big data analytics to detect fraudulent behavior in a cardholder's account before a statement is issued", direct marketing strategies such as the following example mentioned by the authors "Casinos use big data statistical techniques to stream coupons to app users while a gambler is making bets", loyal customer identification and detection of anomalies.

ANN (Artificial neuronal network)

"ANN models are designed to mimic the human brain. Essentially like the brain, which is composed of billions of neurons, ANN models work to make sense of data inputs to make future predictions, reduce informational overload or noise, and classify events. ANN programs do this by simulating what goes on in the human brain by using algorithms to recognize patterns in data and then extrapolate connections, features, and outputs to make decision making faster and more accurate." .The result of this model may be described as the upgraded version of the heuristics used by behavioral economics, which are generalized rules that the brain applies to make hard decisions.

This method can be described as algorithms structured like the human brain, with a large number of examples today as the face recognition developed by the latest devices that thanks to the information input received can develop patterns and provide complex information such as the age of any person, first differentiating the face from the background and then detecting the face composition to be associated with patterns similar as the ones established by our brain to know if somebody is old or young.

Machine Learning

Machine learning consists of gain knowledge from patterns and existing data to provide machine knowledge to form Artificial intelligence (AI) to teach himself.

Machine learning uses adaptive algorithms to build models that clarify connections between relationships, make correlational assumptions, and apply what has been learned to make future predictions.

As more data become available, the machine learns from forecasting successes and failures and then updates predictive algorithms accordingly. (Grable & Lyons , 2018)

The major issue of this statistical method is that the machine requires to be fed with large and high-quality information to provide recommendations that facilitate the decision-making process. "A good example of machine learning is when banks look for fraud as credit cards are being swiped by vendors" (Grable & Lyons , 2018) or as the robot advisors, which use machine learning to provide portfolio management and financial advice to facilitate the investment decision of customers.

The authors also mentioned that the use of these techniques and models are as useful for organizations and governments as controversial, due to the obvious ethical concerns whenever a business or a public authority should be allowed to collect about consumer choices.

Big data risks

In the economic analysis developed by (Tao, et al., 2019) the authors investigate the economic perspective analysis of protecting security and big data, where due to the exponential growth of big data implementation in different areas of society, they could see how "as more consumer and organization information is digitized and collected for data analytics, the potential for cyberthreats and cyberattacks also increases" (Tao, et al., 2019) where understanding the big data phenomenon as linked with the recollection of different information, a large amount of consolidated data can be easily appealing for cybercriminals, especially when in this data can be comprised consumer´s personal and financial data.

In this analyzed paper I found that the large cases and victims of big data security issues were made because the great economic incentives the cyber attackers have when stealing or taking information without information owners consent as the following example "A credit card entry in the black market is sold for about $3–5 in the black market. Personally identifiable information comprising of social security numbers, full names, addresses, passport numbers, and passport images would be sold for as much as $50 per entry" (Tao, et al., 2019).

Thanks to the high profitability of these illegal activities, those malicious actors have several reasons to innovate and employ ingenious methods to successfully attack, which facilitates for this cyber attacker community the availability of more advanced tools at "extremely cheap prices" (Tao, et al., 2019).

Thanks to this two factors (Availability and profitability) the cyber threats became a popular practice that allowed motivated the illegal data practice business, leading to the

appear of new risks , were it became easier to suffer an attack though digital means than it was in the past when attacks involved physical compromises to physical premises, inviting people to take more care of their data with the implementation of mechanisms, security software and special attention when it is about personal data, in order to avoid or decrease those potential risks, but in the case of organizations, special attention should be taken with the implementation of cyber security infrastructures in order to protect the organization´s high value data and their customers data too.

And in order to fix this unfortunate situation "Global forums and institutions focused on the security of cyber-infrastructure have taken to creating rules that every player needs to adhere to so as to uphold the security and privacy of third parties that they interact with, cause regarding the situation from the economic point of view suggested by the author, the cybersecurity risks are costly for the society in general.

As a consequence of this situations the biggest firms around the world had to invest in data security measures that ensure the privacy in their data as a company, but mostly to protect their customers' information to ensure their loyalty, as an example, the text provides the case of the international bank "HSBC who has invested and collaborated with SAS Fraud Management to make use of big data to detect fraudulent behavior in ATM transactions and fraudulent activities. This helps to reduce its global losses from fraudulent transactions and threats, HSBC has started to use SAS Fraud Management as the foundation for its real-time fraud detection and fraud management across its network [29].

The solution is to live in the US, Europe, and Asia, protecting 100% of credit card transactions in real-time. HSBC envisages intensifying the edges to incorporate deception through numerous business and trade networks. Not surprisingly, battling with all forms of

fraud payment cards, online contracts, markets, transactions, and even first-party (customer) fraud has jumped to the top of the corporate agenda [18, 20]. According to Derek Wylde, Head of Group Fraud Risk, Global Security and Fraud Risk for HSBC, they have obtained considerably reduced occurrences of fraud, blackmail, and intellectual property thefts across tens of millions of debit and credit card accounts, suggestively exceeding aggressive objectives". (Tao, et al., 2019)

Big Data: The disruptive phenomenon

*"*Technology is neither good nor bad; nor is it neutral ... technology's interaction with the social ecology is such that technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves*"* (Kranzberg, 1986)

The Big data era is underway, said Boyd and Crawford, regarding how Computer scientist, physicist, economist, mathematicians, political scientist, bio-informaticist, sociologist and other scholars are looking almost obsessively for digital traces left by people such as genetic sequences, social media interactions, health records, phone logs, government records, etc, giving place to ethical discussions about the management of this massive people´s data regarding its costs and potential benefits, triggering utopian and dystopian rhetoric where "On one hand, Big Data is seen as a powerful tool to address various societal ills, offering the potential of new insights into areas as diverse as cancer research, terrorism, and climate change" (Boyd & Crawford, 2012) understanding the power of solving diverse

issues present in the society through the analysis of large data sets that can provide global points of view thanks to pattern construction and automated solutions.

But "On the other, Big Data is seen as a troubling *manifestation of Big Brother*, enabling invasions of privacy, decreased civil freedoms, and increased state and corporate control. As with all socio-technical phenomena, the currents of hope and fear often obscure the more nuanced and subtle shifts that are underway" where privacy limits are hard to define, cause firms try to gather as much as much information as possible about customers in order to have a better understanding of them and customers not concerned about the real value of their data may provide to anybody to ask for it.

With this document is possible to understand the dimension of this phenomenon, that is changing different aspects of our society, resting in the interplay of (Boyd & Crawford, 2012):

Technology:

Maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets.

Analysis:

Drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.

Mythology:

The widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy.

But According to the authors, the most important change that this phenomenon is sparking is in the conception of knowledge as the same as Ford did with the automobile industry, were with the use of specialized machinery and standardized products. "It quickly became the dominant vision of technological progress. 'Fordism' meant automation and assembly lines; for decades onward, this became the orthodoxy of manufacturing: out with skilled craftspeople and slow work, in with a new machine-made era" (Baca, 2004) but it was not only about cars, the different industries understood this new production and business conceptions and applied this tools, giving a large magnitude change to the companies in the world.

This change in knowledge´s definition according to the Boyd and Crawford document can be even better saw in how research is thought today, becoming very close to computational social science, providing "the capacity to collect and analyze data with unprecedent breadth and depth and scale" (Lazer, 2009) leading to a deep discussion about what really matters regarding knowledge, is maybe the objective and undisputed information? Or The subjective information that defend different sciences based in ancient theories since long time ago?

Defining the Big data era, Anderson, Editorin- Chief of wired writes:

This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves. (Anderson, 2008)

Providing a very valuable approach for this research, were the exact sciences with numbers, facts and the possibility to store and analyze large volumes of data are called to provide a better analysis than sciences as the sociology and anthropology, whose vision of the different facts is based on subjective theories used since long time ago, stating that a computer with an advanced software can develop a better analysis if is provided with the right amount of data.

But in the article Critical questions for big data (Boyd & Crawford, 2012), the authors have a different approach, writing : "Taken out of context, big data loses its meaning" an statement that provide a complete vision about big data, that considers the phenomenon´s exact and interpretative dimension, setting that even if a subjective approach cannot provide an exact analysis of certain fact, all the variables cannot fit into a mathematical model, understanding the importance of results provided by exact sciences as math and chemistry but also the knowledge provided by sociologist and doctors interpretation, because what is measured is as important as what and how should be measured.

In order to better understand the previous paragraph (Boyd & Crawford, 2012) cite the following example that illustrate their statements "When mobile phone data suggest that workers spend more time with colleagues than their spouse, this does not necessarily imply

that colleagues are more important than spouses", that's because even if data aims to tie strength through frequency in order to measure the importance of those relationships, "not every connection is equivalent to every other connection and neither does frequency of contact indicate strength of relationship. Further the absence of a connection does not necessarily indicate that a relationship should be made".

Figure 1. Number of people drowned by falling into a swimming-pool

Took from: https://www.tylervigen.com/spurious-correlations
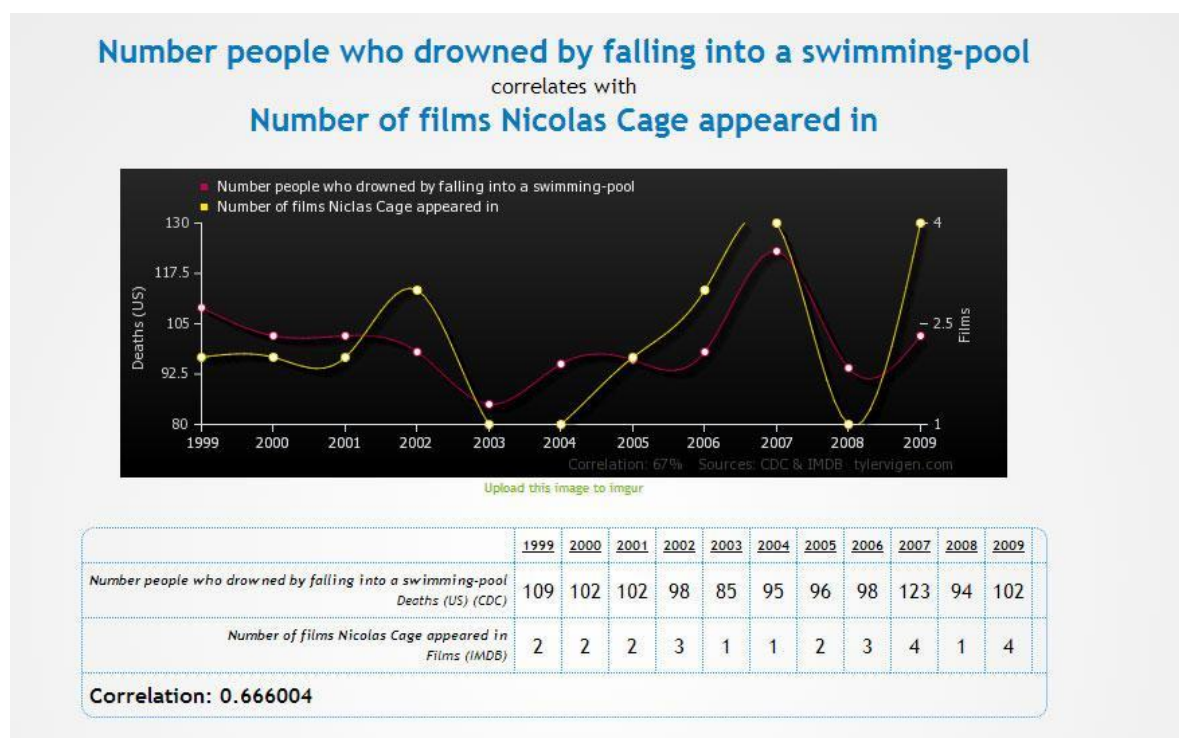
But as this particular situation, when talking about the different approaches that shall be taken into consideration when developing analysis with big data there will be always the risk of practicing apophenia : "seeing patterns where none actually exist, simply because enormous quantities of data can offer connections that radiate in all directions" (Boyd &

Crawford, 2012). In a notable example the website Indy 100 demonstrates that data mining may show strong but spurious correlations between the production of Nicholas Cage films and people drowning in swimming pools in the US, where in average 54.5 people drown in a pool por every Nicholas Cage film.

Main definitions and concept evolution

Since I understood that big data concept has been changing during the years, I decided to take a look to the Ylijoki and Porras research "Perspectives to Definition of Big Data: A Mapping Study and Discussion" in order to compare the definitions of big data and its evolution over the time.

The first approaches to this phenomenon were presented long time ago, when "it was used in the context of visualizing large data sets" (Cox, 1997) and later, associated to data hardware related presentation (Mashey, 1998) and in the data mining context (Weiss, 1998).

Because the different forms of data, researches about different subjects and fields used the big data concept to define large data sets. Cox and Ellsworth use it in the context of visualizing large data sets while recognizing a data breach, stablishing: "Size of data that must be accessed is larger than the size of memory"; later, Big data was associated with Data mining context, where Mashley´s research explain how data management can be provider of valuable information for firms. Those previous mentioned researches are once

of the most important in "the first wave" of big data conceptions, that basically understand it as large data sets with the record of human traces. But in the 2000´s a new definition of this phenomenon gave a totally different approach, providing to big data researches a tridimensional analysis called the 3V´s (Laney, 2001) standing for: Volume, Velocity and Variety.

Dough Laney, as the previous mentioned authors, describes volume as crucial when defining BD but unlike the other authors, for him is just one part of it. He defines Volume as the type and detail of data being collected: "just one par Before the explosion in computing power, businesses and governments collected data but had a challenging time storing what was collected. Today, the volume of data collected from consumers and by agencies continues to grow, but because of computing capacity, storage is no longer an issue. This means that firms and agencies no longer have a data problem but instead have a computing puzzle."

This definition also includes the Velocity dimension, which considers the speed at which data are collected. Data are no longer lag. Instead, data are being collected in real-time at incredibly fast rates and the variety dimension too, that as previously mentioned stands for the different forms of data very diverse such as demographic data, attitudes, and opinions that are captured.

And since Laney´s original research, there have been added more concepts to his definitions, first Veracity, that stands for the possibility to analyze the "Noise" that may be present in data, covering the key attributes that can be provided to the different organizations (Schroeck, 2012) , but as him, some other authors added in their research different variables

to the Laney´s model, as the case of (Membrey, 2013) who took Value into account , because according to him, data only become valuable once is used or processed

2. Methodology

This study has been realized in the aim of showing how the phenomenon of big data changed the relationship between organizations and customers. Understanding how the dig data implementation inside and outside the organization is changing the relationship between those actors considering the disruptive characteristic of this phenomenon through an interpretative approach research.

This observative approach aims to analyze different scientific articles and opinions from different fields and sources to explain the massive impact of the phenomenon contrasting the different theories and studies with their opinion considering the different points of view of the interviewees and respondents.

Having as a reference the different methodologic approaches proposed in (Allard-Poesi, 2015), the interpretative is the most suitable for this document because invites to see the scientific studies as trustful and relevant , considering the diversity of interpretations, rather than absolute reflections of reality(as positivist approach) or entirely fiction who´s content is incorrect (criticism approach).

It is important to notice the importance of the interpretative approach in this document and its special use considering the different interpretations due to the different researches present in this document provide definitions and conceptions in different times and from

different fields, the ethic field may be the most relevant example, that in the case of data managing spark several discussions when considering the relationships between firms and customers.

Qualitative method

This paper is based on qualitative research, often used in the interpretative approach (Allard-Poesi, 2015). The qualitative approach consists in interviewing actors of the research field by formulating relevant questions to lead more or less the interview (Robson, 2002), due to the participants' point of view is necessary to understand the multiple impacts.

The implementation of interviews in this qualitative research allows us to understand the reality of these actors, according to (Patton, 2005)," qualitative research with human beings involves three kinds of data collection: in depth, open ended interviews; direct observations; and written documents". The interviews allow to collect relevant points of view, from people who have experimented the opportunities and risks of big data from their respective fields, making possible a rich contrast with literature

To make the interviews, a semi structured interview method has been used (Robson, 2002),once the most common when doing interviews (Bryman, 2006) "because it has the

advantage of questioning the participants about the subjects of researches while letting them a certain freedom to develop their thoughts on more specific aspects of the subjects" along with questionaries, where trough the same questions I wanted to understand their point of views and approaches when talking about data (CRÉTÉ, 2017), using as an important tool for both methods a questionnaire that ensured that interviewees provide the required information that may allow a detailed analysis.

As Big data became a phenomenon with high impact in our lives, the field of research is particularly too broad because in order to understand its impact in the relationships between organizations and customers, first is important to understand how is changing people´s life in aspects as their other relationships as with friends, family and unknowns and in the case of the company if this phenomenon is changing their way of developing process, to take decisions and manage employees ; that´s why throughout this document the method will be focused in analyze several aspects that are being changed in society thanks to Big data.

3. Introduction

As Big data is a phenomenon that concern different sciences and various aspects of our daily life, several researches have been appearing through the years, explaining their implications for society and also, how firms have been called to the management of this massive data flows.

Even if the use of internet and interactions in social networks have recently sparked the discussions and literature about data in general, scientists and researchers from different sciences have been arguing about the potential benefits and costs of analyzing genetic sequences, consumer behaviors, health records, phone logs, government records and other traces left by people (3), giving place to potential problem solution when being implemented on fields as medicine or military anti-terrorism appliances when using its predictive and analytic power, but also leading to ethical discussions when regarding potential privacy invasions and carefully crafted messages with massive people influence possibility in diverse aspects.

In the beginning, Data sets were obscure and difficult to manage, a reason why they were only of interest to social scientists (3), who used the data gathering and treatment to have a broader vision when analyzing various problems thanks to the relations between the different forms of data, something that made easier cause-effectt studies, regarding the influence of one variable over the others, and used as a tool to have a clearer view of reality, using the data gathering to stablish relations between variables to understand causes and effects whose analysis could be based in facts rather than just interpretations.

But as the Big data applications were called to provide smarter solutions, the required data became too large to be analyzed and the variables too different to be included in a same model, inviting researchers and data scientists to develop new gathering methods and processing information alternatives that facilitate the processes and allow the large data variety to fit in the same analysis and to be pertinent when solving the same problematic, developing more efficient methods during the years.

But due to the data analysis was called to provide more complex information, the researchers required automated processes to ensure the accuracy and to make possible the recollection of large amounts of data. That´s why, in 1890, the US Bureau of census gave birth to the punch card machine, the world´s first automated processing equipment, allowing that once collected, the data could be classified and compared in a faster way, delegating the mechanic processes to machines and allowing the sciences to focus in the results interpretation.

But as social sciences where not the only interested in understanding phenomenon and facts in their field, Physicist, economics, political scientists, bio informatics and many other sciences

decided to implement data analysis (databased methods) to adopt a quantitative approach in their systematic processes in order to support their researches in facts.

With all the professionals we mentioned before conscious about the benefits that data can provide to their jobs and daily life and interested in those data possibilities it became easier to understand how the world is turning around big data with a society that has been taking profit from this massive data era but also beginning to question the possible treats that may come with this big data focused world. In order to understand the previous statement we should consider our interactions in social media, were the geographic distances are no longer a barrier, facilitating the communication around the world for everybody making easier to keep strong distance relationships during the time and enabling business to develop internationalization strategies.

All of this are once of the most important advantages of posting our data on those new digital platforms as the social media, but as we previously mentioned, several studies and the recent spreading of data sciences have proved the potential threats of posting our personal data in sites as the different social media or everywhere online, because as our information is registered in more databases our risks to be privacy invaded or that our information can be sold may be increased.

But even if we have to be concerned about what is being done with our personal information, it doesn´t mean that data is harmful for us, throughout this text we will analyze the several opportunities that data management is providing to the different organizations, developing better relationships with customers and all their stakeholders that contributing with efficiency to the society will reach more benefits in general.

4. Today´s influence of Big data in customers and organizations

4.1 Customer

With the development of a large variety of technologies and techniques to identify, measure and influence people. Smart devices such as smartphones and wearables assists today's users in different aspects of their life (Kappler, et al., 2017), arousing practices that bear several risks for individual rights, that even if foster societies self control as in the case of china, were through a face recognition system, citizens behavior can be constantly regulated, is also seen to profoundly change the healthcare systems and the treatment of disease (Kappler, et al., 2017)

Self-tracking

"By means of smartphones and wearables (Smart watches activity trackers, smart glasses, etc.)individual behavior and bodily functions are now accessible for informatization too" (Kappler, et al., 2017). This, often with self-reflection objectives, lead people to track daily activities with aims to obtain social recognition, but also looking for the possibility of curing diseases, controlling emotions, anticipating risks or increasing performance through the instant feedback provided by those previously mentioned devices.

From sociological perspective, self-tracking has been instructive as it allows to study the impact of information in empirical fields, allowing a more detailed study of events trough constant quantification that permit the measurement of progress and the accomplishment of

already set objectives in order to solve problematics and understand various phenomenon themselves and their possible impacts.

The Fact is that human being have always been quantifying, Actually, important figures in history attribute their success to quantifying and tracking different events and happenings in their life, with the objective of sharing ideas, interpretations, and points of view and also for experimenting, two of the most important uses of data that have been present since long time ago in our society.

"Benjamin Franklin, the eighteenth century statesman, kept account of how he spent his time and whether he lived up to their virtues he set forth himself" (Neff & Nafus, 2016),in order to have a tracking as a plan examination, actually in eighteen century people used diaries as a way to share small pieces of knowledge, register stories and personal logs. Also this practices were made by inventors like Richard Buckminster Fulle who created a giant scrapbook of sorts, were he recorded ideas archiving it with high frequency.

The last paragraph shows how old is the record of massive information, and how it has been done several years after the implementation of internet and social networks. The scholar writer Lee Humphreys found that "Twitter feeds bear a great resemblance to eighteenth and nineteenth century diaries in how both "Account, reflect, communicate and share with others using media of the times" (Humphreys, 2018), Showing the involvement of human being in different activities related record daily traces and share personal data and information to other people.

Aside from the communication and sharing uses, data was also helpful in early years in experiments, conducting to the discovery of malaria vaccines, ibuprofen and the

understanding of neuroscience but in those cases rather than aiming for information sharing, the scientists tried to use the data as source of answers. As an example we have one of the most popular studies made by Sir Isaac Newton who nearly blinded his eyes trying to understand the humans eye working with sun reflection analysis.

What cannot be measured, cannot be improved

Record different sort of traces have been for humans an important activity that can lead to obtain better quality information or to track actions while looking for a big goal, and as we saw in the last paragraphs when the data was short and easy to register by hand, how much a person could write or how much of this data could be measured was the limit, but nowadays, the big data phenomenon changed how the traces were seen and managed.

To explain the strong link between the concepts of Big data and Self-quantification, the Fitbit example seems to be the most accurate. Fitbit is a technology organization created in California, whose most popular product is the Fitbit watch in it´s different series. The device collect real time raw data related with users daily activities, training, nutrition, weight and sleep quality such as number of steps per day, heart rate, burned calories, user weight and hours slept per day, in order to transform it into useful insights that provide a clear view of daily behaviors impact in user´s life.

To conceptualize the link between both concepts is important to take into account once of the most relevant definitions as Dough Laney´s concerning Big data, setting that this concept rests in the interplay of the 3V´s: Volume, Velocity and Variety (Laney, 2001):

Figure 2. Dough Laney´s Big Data Key concepts
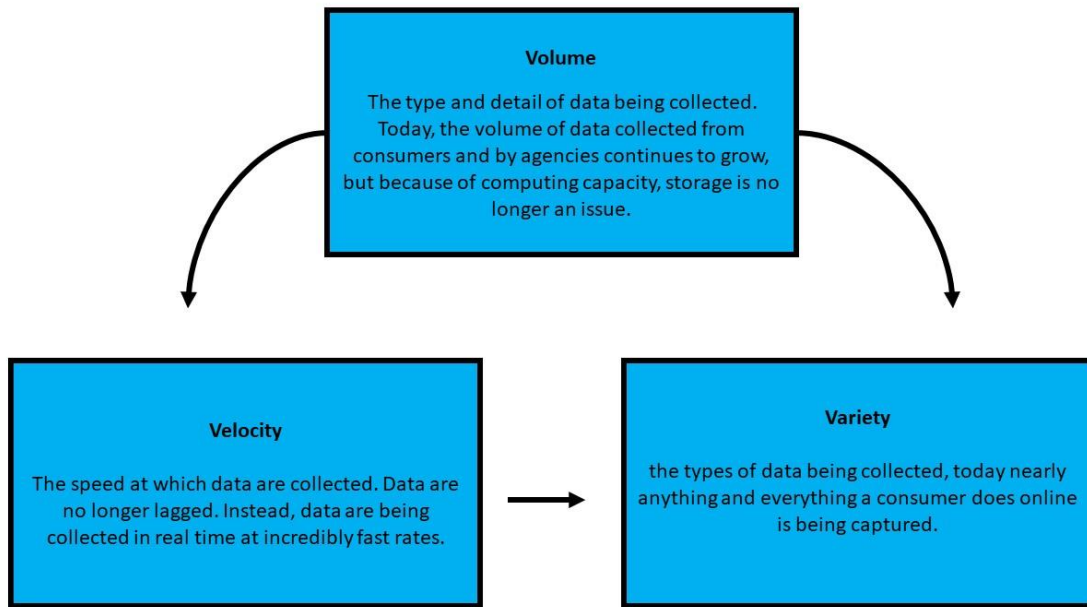
Source: (Laney, 2001)

Those dimensions can be also found in the Fitbit watches:

1. **Volume:** Fitbit stores high volumes of data (for example number of steps in a day) with almost not limits, thanks to the large storing capacity of the device, but mostly, to the cloud computing options were the data is stored.

2. **Velocity:** Fitbit gather real time data in order to show relevant information related to users current situation and develop meaningful insights.

3. **Variety:** Fitbit requires different sorts of data like number of steps per day, heart rate, burned calories, user weight and hours slept per day, that permits the current health reports and recommendations that lead to the goals achievement.

Security Issues

The different Big data applications are giving to personal data an additional value, due to individuals and organizations identify in this type of data the opportunity to see in each person´s data, information that would determine their behavior and group people with similar characteristics, allowing that according to Boyd and Crawford big data present a big brother manifestation that made always easier the privacy invasions, civil freedoms decrease and an increased state and corporate control.

Letting people and firms to have access to sensitive data, like the maximum amount of money that somebody is willing to pay for a product, or as the multiple cases of insurance organization big data implementation, was their risk can be diminished if they are able to know if anybody is developing an illness or specifically know which of the assurance package services the person is going to use the most, in order to set the price.

As I founded data security as once of the most important subjects of this research because the large risks and information that is important to know in this days ,when our privacy is such a valuable asset, I decided to interview a professional specialized in technological management and security with large knowledge about the different dimensions when talking about data security.

Mr. Yves Barlette is a Professor of Montpellier Business school with a PhD in Management Information Systems at the University of Montpellier and is a current "member of the MRM research centre, being focused on actors' behavior related to information

security, publishing articles in peer-reviewed journals such *as* International Journal of Information Management *(IJIM),* Systèmes d'Information et Management *(SIM),* Production Planning and Control (PPC)*, Journal of Global Information Management *(JGIM*) and several books and book chapters."(Montpellier Business School, 2020)

Before the interview with Mr. Barlette there was a large questionnaire prepared in order to understand the different issues present in the information system field and he´s personal opinion about the Big data panorama , as the first questions were mainly orientated to understand the main reasons for organizations to gather data, he was asked with ¿why does firms were looking everyday for more and more private information ?,answering that more than looking for deeper information, they acquired the capacity of store large amounts of data, becoming almost limitless in customer information storage, thanks to the reduction in costs that services as cloud computing were offered to organizations. Allowing the firms to register in their databases diverse information about what we do, how and when limiting our privacy more than ever.

" 20 years ago, an average person was registered in approximately 500 databases, today can be millions"

Is a truth that individual firms never asked us to explicit fill large formularies with our private data, but the quote previously presented shows how our information is there, handled by firms that can use them for different purposes and in the end can be shared and complemented between them, establishing strong connections that can lead in a better understanding of customers. In the interview, Mr. Barlette mentioned the example of

Facebook, a firm that has access to different kinds of people´s information, and even if we often saw a data privacy policy they mention that our data can be only shared with few associates, but at the end of the day, the biggest firms are their associates.

As an example of the organizations with Facebook shares users data with, is necessary to cite the Cambridge Analytica case, were after the USA presidential elections in 2016 and the Brexit, several probes shown that citizens votes were influenced by carefully crafted messages created by politicians and data Scientifics of Cambridge Analytica with the people´s data stored by Facebook, allowing massive access to Facebook´s users' needs and fears, using private and sensible as their stronger weapon to win in the Donald Trump´s presidential race and the United kingdom exit from the European Union.

When asking Mr. Barlette about his perspective about the future of a world with the possibility of personal crafted messages, he said that there were important opportunities for people and firms as the possibility of avoiding thousands of irrelevant advertisements and instead, the exposition to just a few that really matters for the receptor and aims for solving a real problem for him and for the side of the firm, those will be messages with higher response rates from the objective; but also it may spark threats for people and society in general, such as the risk of influence and manipulation, causing that people take actions based in disinformation and particular interests.

4.2 Organizations

As we could see in the previous paragraphs, people have found in the collection and analysis of data the way to understand and record phenomena and problems that have

occurred throughout history. Similarly, companies have been concerned about making data a tool that allows them to achieve efficiency and effectiveness in each of their actions and processes, achieving new capabilities as the following:

Global vision

As the firms began to have access to different sorts of data in large quantities, their vision of the business became more wide and complete, externally, the firms acquired the capacity to read the environment with more accuracy, thanks to a better understanding of stakeholders through data analytics that allowed a more detailed study of their required inputs and also how their outputs could beneficiate the organization.

One of the most important consequences of the global vision is the massive international expansion of firms that became possible thanks to the easy access to information about the different markets around the world regardless of the physical distance and rather focusing in the operations and business opportunities there.

In the other hand, this global vision also helped to achieve a better performance inside the organization, in the way that allowed to understand and register every single process developed, making difference between what is well made and what is need to be improved and the same happened with the employees, who thanks to the data implementation could develop tasks according to their skills while working on their weakness in order to achieve better performance.

Decision-making improvement

The implementation of massive data analytics allowed to improve their decision-making process, in the way that much more variables can be considered ensuring multiple perspectives of different areas in the organization before taking a significant decision, but to ensure integral decisions, the firms must to provide data access to everybody inside the firm and not only to be stored and owned by the IT department. This information democratization even if can be considered as too bureaucratic will provide more complete solutions to the head quarters with reliable back ups based in objective information rather than intuition and gut.

In order to better understand how the democratization of information can help in decision making ,Walmart (the world's biggest retailer with over 20,000 stores in 28 countries according to Forbes (Marr, 2017)), provides a great example with their Walmart's data café "a cutting-edge analytics hub, where hundreds of streams of internal and external data can be sliced and diced to deliver insights (Kappler, et al., 2017)",Simplifying the large amounts of data into relevant information that can be more easily shared and take action with.

Predictive power

Once analyzed the main tools that data offers to firms is possible to understand how data management became essential for any kind of organization no matter the industry even if some are more clear in their use because recording and measuring are allowing organizations to understand the business and take better actions in the turbulent environments where they are developing.

With the aim to take a closer view to data management in the organizations, I decided to analyze how knowledge conception has been transformed with big data through the study of one specific organization in order to reach a deeper and detailed understanding about the effect of this phenomenon in the people and daily tasks developed in the organization.

5. Organizational Analysis

FAM Maritime Agency

Is a logistics solution company composed by several maritime transport agencies such as CMA CGM, ULS, and Overseas, that provide cargo shipping transportation around the world with different services, addressed to different clients and have different perspectives of the business but whose core is the achieving the operations efficiency, adaptability and personalized assistance through the implementation of leading technology. As I wanted to understand how big data is changing business I decided to focus on their people data perception (See appendix for the answer and question transcription table), interviewing employees in different areas of the organization in the following charges:

Customer service coordinator: Karla Patiño

National operation director: Victor Plazas

CHRO (Chief Human resources Officer): Paola Yara

Commercial manager: Sergio Ortiz

When interviewees were asked about the forms of data present in their workplace, as they had different information to manage and different tasks assigned, their answers shown different data conceptions, form raw data: "Weight, loading and unloading ports imported and exported products" as the Customer service coordinator mentioned, to HR software "that consolidate the most relevant information about workers" concerning their performance and contract information as the CHRO mentioned. Data that according to the interviewees had to be protected because they considered costly and valuable for competitors.

But as the organization is operating with more and more data I considered important to know how is the firm and the employees preparing upgrades in the threat of raw data, were   Ms. Patiño mentioned that thanks to the frequent software updates made by the organization is becoming easier to find synthetized information to be used when required and as the commercial manager said "The way to treat data have been changing too much", due to the core of the organization is the implementation of technology to provide the best solutions.

As we discussed before, once of the biggest challenges of data implementation is that firm members know what to analyze from data, that´s why regarding the long career of the interviewees I decided to know how do they manage the mix between data and career experience (through the question ¿how you think your knowledge and experience help you to read the data present in your daily activities?) where I got interesting answers such as the one gave by the commercial manager: "With experience, it is already easier to identify the

key points of the information we use to use it for the benefit of our business, which allows us to arrive with a service more in line with the needs of each client" leading me to understand that a reason why a large experience employee may earn more money than a new one maybe not because the amount of data that can be mined, but the quality of information can be taken from this data and also the speed which it may be done.

But what if we do not compare a Highly experienced employee with an inexperienced one , but instead with an advanced algorithm machine, a discussion that has been under the table since some years ago, and our interviewees also were invited to participate on it, arguing in the case of the CHRO that "there are factors inherent to people that must be combined", defending the need of feelings and human´s natural judgment to make decisions and develop actions, similarly as the answer of the commercial manager, who said that even if a machine can talk to clients, the communication won't be the same.

That´s because hopefully (I think) the employees of this organization are clear that their job goes beyond the interpretation of data because in their workplace they find that this is not enough to accomplish their different goals, are the cases of Mr. Victor, the national operations coordinator who said that due to several circumstances that arise in shipping, there´s need of human interaction "to respond for all unforeseen events that usually occur" and in the case of Human resources department Ms. Paola who also argues that direct contact with people and support are essential when treating with people.

With all of this, we cannot say that data and treating with people are opposite sides, actually, managing correctly the data can lead to better treatment of customers and the following answers of the interviewees are a probe of this assumption: "Relationships are based on the acknowledge you have about the counterpart, so the better you know your client,

the best relationship you will build" were Ms. Karla holds that a better understanding of clients can lead to a better understanding of them; "Yes, we can see important trends and also try and anticipate to shifts and behavior", Mr. Victor on his side find it as a tool to anticipate interactions; "Yes, because under the understanding of their needs, better services can be offered" on her establishment, Ms. Paola also find in data a tool to better satisfy the customer needs.

But in the interview I did not consider customers as the only personal treats with they deal, so, with the question ¿Do you think that having more information about your office partners, you can build better relationships with them? I wanted to know how data and patterns stored on their minds could help with their colleagues' daily relationship, were Mr. Victor´ s and Mr. Sergio with their respective answers "Yes, understanding backgrounds can tell you a lot about a person's motivations" and "Yes, because the communication can be improving and have a better result on the daily work" find in their colleagues' data understanding a way to achieve efficiency in teamwork but Ms. Paola arguing that " knowing people helps to have better communication, because of their communication preferences and ease of interaction you can find the best way to reach them" understand colleagues personal data treatment as a way to improve communication with them.

I found in their answers about this specific question, the sensing characteristic of data, that in the interview with Mr. Ives Barlette he defined as the ability to read the environment in order to have a more clear and accurate view, but in this same interview with the MBS professor we were told that firms need also to implement the responding characteristic of data, that will allow the firm to develop quicker and smarter actions than the competitors in order to "always be ahead", but these actions are just the consequence of the right decision-

making process, which I wanted to know how dependent was from data with the question: ¿Why do you think different sorts of data can lead you to better decisions? Where the CHRO mentioned that "Yes, they are definitely necessary …to determine the action plan to strengthen the culture and prepare employees to improve the performance of their functions. Because data, sets a reference and this allows to take the necessary actions to approach the purpose", showing the power of data in the different actions developed by the firm.

The analysis of these interviews let us understand how important it is becoming the data management no matter the business or the industry because of the unlimited sensing and responding possibilities that it offers for diverse purposes. As an example we have Amazon, one of the biggest retail online stores in the world that has been pointed out because practicing price discrimination, it means that two consumers might be paying different prices for the same product, because Amazon know information about them like how much they need it and how much are they willing to pay.

This may sound unfair, but imagine to have a tool that may adjust offerings for clients who can´t pay for the whole product or service, or to identify the clients who can pay a little more, and in the case of our interviewees, they all agreed that their negotiation power would increase with such a tool, and in their answers, they included advantages as the possibility to be adjusted to the specific customer needs and the implementation of cost-saving strategies implementation with high possibilities of maximizing earnings.

6. Results

6.1 Data have been since several years having an impact in humans society

Humans have been quantifying since a very long time ago, using data for understanding diverse events and facts in their daily life finding appropriate solutions. It means that data applications used in this research are not entirely new and that even if hasn´t been used in advanced computers with multivariable analysis have been used as tools in several appliances as information gathering through different opinions, looking for people´s behavior understanding and pattern associations with previous experiences when testing new processes aiming for efficiency.

6.2 Society is having access to better products and services thanks to organizations big data

6.3 Implementation

Today´s organizations are constantly storing different sorts of data, but it doesn´t mean that their uses are only to harm society, and even if there are some risks that everybody should be concerned about as the potential privacy invasion, aggressive price discrimination and massive manipulation as we saw in the analysis, the organizations in the different industries are implementing big data solutions looking for efficiency improvement thanks to the sensing and reacting possibilities offered on its diverse appliances, allowing a faster and meaningful transformation of raw data and a better understanding of the whole environment, permitting to take better decisions thanks to the availability of tools and the information democratization for employees in the different areas, who can have a holistic vision of the business, and with this also a better understanding of demand needs, improving the efficiency

inside organizations and ensuring better offerings to customers innovating with better solutions to the market.

6.4 Impact of Big data in knowledge conception

Since big data, the knowledge conception in firms and society changed too much, and as our approach aimed to understand this phenomenon from the point of the different science of view, it was possible to evidence their analysis methods. In the beginning, the knowledge was based on empirical and theoretical statements but nowadays, the application of computational and statistical sciences are providing a more detailed analysis for every single situation with stronger bases founded in multivariable analysis and complex algorithms.

7. Recommendations

- Every day new technologies are available for organizations, that´s why the firms shall be always informed of the big data new trends that may improve the business.
- Cyber attacks are becoming more frequent every day because of the various and cheaply available tools for hackers, who are constantly looking for new opportunities, due to the huge profitability of their attacks. That´s why organizations and their IT departments shall be very attentive

- Today there are several ethical discussions about the uses of customer´s data by the firms, raising more conscious consumers about the uses of their data, that's why organizations shall be responsible with their customer´s data.

- Organizational environment is constantly changing, but even more when we talk about technology implementation, that's why organizations shall implement the data analysis tools in order to be updated about trends, possible threats, and opportunities

- As we mentioned before, decision making is once of the most important processes inside the organizations, that´s why firms shall invest in big data gathering and analysis tools in order to base their decision making in different sorts of data developing a more complete analysis of situations.

8. Conclusion

Thanks to this document we could evidence different perspectives about the impact of this phenomenon, considered disruptive because of its various implications that we previously defined in different aspects in our society revolutionizing not only the way organizations interact int with customers but also the way we develop our daily tasks, define the knowledge and communicate with other people and our environment, focusing in the role of organizations as protagonists of Big data implementation giving place to ethical discussions due to potential privacy invasions, but mostly because people´s life improvement through better offerings thanks to the different technology uses.

9. References

Anderson, C., 2008. The end of theory, will the data deluge makes the scientific method obsolete?. *Edge.*

Baca, G., 2004. 'Legends of Fordism: between myth, history, and foregone conclusions. *Social analysis,* 48(3), pp. 169-178.

Boyd, D. & Crawford, K., 2012. CRITICAL QUESTIONS FOR BIG DATA. *Information, Communication & Society,* Volume 15:5, pp. 662-679.

Cox, M. E. D., 1997. Application-controlled demand paging for out-of-core visualization. *Proceedings of the 8th conference on Visualization,* p. 235–ff. .

D., P. A. A. L. A. S. B. A. B. D. C., 2009. Computational social science. *Science,* 323(5915), pp. 721-723.

Grable, J. & Lyons , A., 2018. An introduction to big data. *JOURNAL OF FINANCIAL SERVICE PROFESSIONALS,* 72(5), pp. 17-20.

Humphreys, L., 2018. *The qualified self: Social media and the accounting of everyday life.* 1 ed. s.l.:MIT press.

Kappler, K., Schrape, J., Ulbricht, L. & Weyer, J., 2017. Societal Implications of Big Data. *Springer-Verlag GmbH Germany,* pp. 1-6.

Kranzberg, M., 1986. Technology and history: kranzberg's laws. *Technology and culture,* 27(3), pp. 544-560.

Laney, D., 2001. 3D data management: Controlling data volume, velocity and variety. *META Group Research Note,* 6(70).

Marr, B., 2017. *Forbes.* [Online]
Available at: https://www.forbes.com/sites/bernardmarr/2017/01/23/really-big-data-at-walmart-real-time-insights-from-their-40-petabyte-data-cloud/#617c5266c105

Mashey, J., 1998. Big Data... and the Next Wave of InfraStress.

Membrey, P. C. K. D. Y., 2013. A disk based stream oriented approach for storing big data.. *Collaboration Technologies and Systems (CTS),* p. 56–64. .

Vigen Tyler, . (2019). *Spurious correlations*. Took on november 20 of 2019, de Tylervigen website: *https://www.tylervigen.com/spurious-correlations*

Neff, G. & Nafus, D., 2016. *Self-Tracking.* s.l.:The MIT press.

Schroeck, M. S. R. S. J. R.-M. D. T. P., 2012. Analytics: The real-world use of big data, IBM Global Business Services, Somers. *IBM.*

Tao, H. et al., 2019. Economic perspective analysis of protecting big data security and privacy. *Elsevier,* Volume 98, pp. 670-671.

Weiss, S. I. N., 1998. Predictive data mining: a practical guide. *Morgan Kaufmann. .*