South Asian Journal of Engineering and Technology Research Article

Open Access Full Text Article 1(2020)1-6 DOI: 10.26524/sajet.2020.5

Big database technologies: shaping the future world

John .M[a]
[a]University of Maiduguri,Maiduguri, Nigeria

Corresponding author.
Correspondence: John .M
E-mail:matthewjonny086@gmail.com

Article info
Received 7th May 2020
Received in revised form
14 June 2020 Accepted
17 June 2020

Keywords
Double-digit growing market,
increasing interest

Abstract
Exponentially growing datasphere, double-digit growing market, increasing interest, many successful projects. When big data is joined by analytics, everything is possible. Big data analytics are the protagonists of the IT market in a massive way in all sectors. In the coming years, the job market will require many experts in this sector, and many professions will be transformed. But with what impact on data protection? Will GDPR help balance technology and privacy?

1.Introduction

Big data has long been one of the hottest topics on the IT market, and beyond. These are data sets so large as volume, and of such complexity, that they cannot be managed using traditional software tools. A classic database, in fact, manages the data in rows and columns, but the big data data cannot be so easily pigeonholed, because they are not available in a structured form: they can be present as documents, values detected by IoT sensors, meta data, geographic locations, and so on. Big data are not "big" data only for their size and variety, but also for how important they can be for the companies and entities that analyze them. Thanks to the use of increasingly sophisticated tools like artificial intelligence, companies are able to find hidden gems within them, and thus obtain the information necessary to make more informed decisions [1-4]. According to IDC, data represent the new basis of competitive advantage, whether structured or unstructured, generated by man or machines, stored in the data center or in the cloud. By exploiting large quantities and diversity of data to discover patterns and pursue revolutionary ideas, a company can win the war in the growing competitive landscape of companies. Storage is an integral part of an organization's data strategy as it actively contributes to the process of archiving and analyzing information. For IDC, the challenge is to build storage systems that can handle such large volumes of data, but keep costs low, without compromising performance [5-8].

2.CONTINUOUSLY GROWING DATA

The volume and variety of data continues to pervade organizations at all levels at an ever-increasing rate, convincing managers to derive value and determine its impact on business. All our daily activities produce data: user posts on Facebook, a search on Google, the use of an app, the purchase on an online site, but also in physical stores, with credit cards and loyalty cards. And again: a photo, a video, a voice message, a tweet, the itineraries, the comments, the "likes". Not to mention the data created by the objects interconnected to the network: just think of the intelligent infrastructures of cities, the sensors mounted on buildings and means of transport, and smart appliances. Each of us constantly contributes to data production. To give an idea of what the term "we" means, we must take into account the number of people who make it up: there are included Internet users, who in the world are almost four billion, those who are active on social media, almost three billions, and the owners of smartphones, according to various estimates over 4 billion, these numbers are constantly updated because they are constantly growing. It is clear, therefore, that with these numbers the amount of information collected in real time is exponentially expanding [5].

3.THE MARKET SCENARIO

The interest in big data is remarkable, and it is not measured only in words, but also with the huge investments of companies in increasingly powerful computers, increasingly capacious storage and increasingly sophisticated algorithms: thanks to them, big data becomes a great opportunity to extract useful information, revealing connections, to create detailed profiles about us, unique user experiences and personalized commercial opportunities. In the IDC classification, the big data and analytics (BDA) software market actually includes several areas that functionally solve all the phases of the life cycle of decision-making and decision automation processes. These processes are often referred to as business intelligence, business analytics, analytics, big data, and so on. The BDA software market consists of three main segments: analytics and performance management applications (APM app), business intelligence and analytics tools (BIA tools), analytical data management and integration platform (ADMI platform).

4.THE DATA MONETIZATION

As we said, companies have understood how data represents real wealth, still underused, or even, in some cases, poorly managed. In-depth data analysis operations, however, can help improve the main business functions, and can also give rise to new business opportunities. In every market sector, the BDA can create benefits: the attention to the data is transversal to all sectors, in the public and private sectors, in production companies and in suppliers of services and utilities: everywhere, big data represent potential huge, because within the amount of data there is very important information, where the importance, of course, is linked to what analyzes are made. The ability to transform data into value, not only theoretical, but precisely in terms of turnover, is called data monetization: it is a process through which corporate organizations take the data in their possession and that generated by the interactions with their business partners, customers and suppliers, internally [7].

5.BIG DATA IN FINANCE AND RETAIL

According to the latest report of the Polytechnic Observatory on big data, in Italy the banking sector continues to represent the one where there is the greatest use of these technologies (28%), followed by manufacturing (24%) and, at a distance, by telco & media (14%), other services (8%), PA and healthcare (7%), large-scale retail and retail (7%), utilities (6%) and insurance (6%). The sectors that grow the most are insurance, manufacturing and services. The banking sector is particularly data-intensive: here huge volumes of information must be managed, which often prove to be a real treasure of data from which valuable insights on customer behavior can be extracted. Among the various objectives, the main ones are understanding and satisfying the customer, in the various activities, and checking compliance with regulations. Particular attention is given to the prevention of fraud: every year there are losses caused by fraud and various crimes, ranging from cyber attacks to money laundering. With analytics, you can find solutions that allow you to take proactive and comprehensive approaches to fight fraud [8].

Another specific feature of the sector is risk analysis: for a bank, having updated data on portfolio risk allows risk managers to monitor indicators and constantly align front office activities with the risk appetite policy of the bank. In the GDO and retail sector, the relationship with customers is fundamental, and the best way to manage and optimize it is through big data: it can perhaps be said that this is probably the sector where we first understood what wealth was hidden in the data collected every day from customer transactions in physical and online stores and, in recent years, also in forums, blogs, online magazines and social networks. Retailers with big data analytics are able to better manage their customers, make transactions more effective and understand what actions to take to keep the volume of business at an adequate level. Thanks to the analysis on big data, it is possible to optimize and personalize its offer on the habits and preferences of each individual customer, defining for each one a precise profiling, identifying

which new products or services to provide, and monitoring the opinions and the level of satisfaction of the customers to improve the shopping experience, develop cross-selling and up-selling, reduce the risk of abandonment and attract new prospects. Analytics are also used to increase efficiency in each phase of the supply chain, from purchases to positioning the goods on the shelf. The data, properly processed to make them anonymous, can be sold to suppliers: in this case, they could be useful to understand how much their products are liked, thus facilitating targeted promotional strategies [6-9].

6.BIG DATA IN TELCO AND UTILITY

In the telco sector, companies operate in a ruthless environment, with strong and continuous pressure on prices, with volatile consumers who are ready to take advantage of the various offers that appear on the market almost every day. Telephone operators have always collected large amounts of data on a daily basis, often used only for charging or for network maintenance activities. This gold mine, only for a few years, has been used to exploit its contents, crossing traffic data with data related to network disturbances can be useful for improving service levels. Probably, the top priority for sector managers, thanks to the analysis of big data, is to be able to create a unique experience to increase consumer involvement: from customers' preferences and behaviors, from the use of devices, from demographics , from their feelings and opinions, they can produce and deliver personalized content anywhere, anytime and on any device. The aggregation and analysis of data from various sources, including billing, social listening and use of plans, can be used to predict customers at risk, and proactively offer an interesting loyalty offer. The simulators and forecasting models can estimate customer acquisition and loyalty, and can be used to allocate advertising costs [10,11].

7.MANUFACTURING AND INDUSTRY 4.0

In the manufacturing sector, companies analyze data to optimize their production process. The goal is to improve production times and processing cycles, reduce waste, dead times, waste, errors in production and cut costs. All with an eye on quality. The BDA is also important for optimizing its supply chain, in terms of times, costs and supply methods. For a manufacturing company, the maintenance costs of production machines and related plants are often very high: it is therefore important to know the life cycle of the equipment in order to plan maintenance and component replacement interventions. It is also crucial to analyze the machine data to quickly identify the causes of the faults, not only for the repair phase, but also to modify the design, if this serves to reduce the faults themselves [10-15].

8.SUCCESSFUL PROJECTS

We see below some successful projects, developed in different sectors, mostly in large companies. We thank SAS, IDC (www.idc.com/italy) and researchers from the University of Pisa Paolo Cintia and Luca Pappalardo for their collaboration.

Enel Green Power - In the utility sector and in particular renewable energies, the use of analytical models to identify areas of malfunction within geographically very large plants offers significant business returns. The use of big data in this sector has made it possible to highlight and resolve malfunctions in a targeted manner, improving the production efficiency of the plants and reducing recovery costs in the event of failure. In addition to identifying and resolving malfunctions in an immediate and timely manner, with analytics, Enel Green has been able to increase production efficiency and ensure the operational continuity of the plants; reduce the costs of restoring production plants in case of failure; understand and solve long-open problems.

Sky - Thanks to analytics and customer intelligence, Sky has managed to reach the single customer with the most suitable offer, through the preferred contact channel, without ever losing sight of budgets and contact policy constraints. Associating the benefits for the consumer with the objectives of the company means

allowing consumers to choose based on a satisfactory price-quality proposal and in line with their preferences. The project consists of two souls: an analytical component to construct predictive measures aimed at estimating future customer behavior; an operational component that, through the optimizer, associates individual offers to each customer for the chosen optimization period. In this way, Sky managed to achieve incremental redemptions, ranging between 10 and 20 percent.

Vodafone - Thanks to the omnichannel analysis, Vodafone has defined the stages of the customer journey to optimize the customer experience on all touch points. The most frequent use of big data by Vodafone is aimed at acquiring greater knowledge of customer needs to improve the customer experience, make the commercial strategy more effective and optimize the performance of the networks in order to provide a service. excellent. Big data are also used for real time marketing, offering customers what they need most, increasing their trust in the brand, and therefore optimizing value for the company. Another area of application is that aimed at improving the network: thanks to direct customer data and reports, temporary outages can be intercepted faster, thus intervening to improve networks.

Octo Telematics - Octo Telematics has data enhancement at the heart of its strategy. The main intent is to create new knowledge by exploiting a multiplicity of data to obtain information capable of generating value for companies and policyholders. Octo Telematics has an analytical platform that uses a set of data to reconstruct the customer profile, opening new horizons in terms of packaging the offer. The success story of Octo Telematics is an example of IoT technology and analytics implementation. A device installed on the cars transmits the data detected by the sensors in real time, allowing mutual recognition between the company and the customer

Leroy Merlin - Leroy Merlin's business need is to reinvent the purchase path in a cross channel perspective to optimize interaction with customers. Faced with the need to have available a system capable of automating forecasting and ordering activities for individual operators, Leroy Merlin has modulated the parameters based on the result. In this way, the main results achieved thanks to analytics include: reduction of delays and optimization of the entire inventory chain; improvement of business performance and customer satisfaction level.

CNH Industrial - CNH Industrial, among the world's largest capital goods manufacturers, has built a data management system to collect and analyze warranty data and information on the use of products. The goal is to measure quality KPIs to continuously improve the dependability (reliability, robustness, durability) of the goods. Thanks to analytics, CNH manages to obtain crucial information to determine the quality of the goods and to identify any component defects:

Disney - Service companies are heavily focused on integrating analytics at every stage of their customers' travels, while maximizing profits and revenue. Disney's efforts in this area show the impact that big data analysis can have on virtually all areas of a service organization. As part of the customer experience, Disney is focused on creating an immersive, perfect and personalized experience for its guests. For example, FastPass + allows guests to book travel times, restaurants and shows in advance via a website. To understand and target customers,

UPS - Being a constantly moving company, UPS stores large volumes of data, most of which derive from sensors placed on their vehicles. These data not only monitor daily performance, but have also caused an important metamorphosis in the structure of UPS driver routes. The initiative was called ORION (On-Road Integration Optimization and Navigation), and is probably the largest cooperative research project in the world. It relies on the data provided by online maps to reconfigure a driver's pickups and deliveries in real time. The project aims to save 32 million liters of fuel by cutting 137 million kilometers from daily routes.

9.THE FUTURE DATA MANAGERS

Even according to the majority of respondents to the Hays group's "Jobs of the Future 2018" survey, published last February, digitization will not make traditional professions disappear, but workers in all sectors will have to acquire IT skills in order to remain competitive. on the market. Among the professionals interviewed, one in three declares to update already every six months, another third even every three months. Industry 4.0 will lead to the requalification of professional skills, and companies will have to invest heavily in training, seen as a strategic corporate asset not to risk losing competitiveness. The three professions with the highest growth in job opportunities in the next 12/24 months will be the cybersecurity specialist, responsible for digital security, the data analyst, and the growth hacker, a professional who deals with growth through the use of advanced lean marketing and software. Mirna Packages of InTribe details the tasks of the data analyst: «The data analyst is the operative, very often with a degree in statistics, capable of analyzing large amounts of data with the aid of software, transforming them into information with high added value for the company. Its main goal is to facilitate decision making and problem solving. There are several variations of this professionalism: the predictive analysis expert, who is able to analyze the data in order to prevent machine breakdowns. The emerging roles of artificial intelligence is reshaping the world for better society [16-18].

Conclusion

When newspapers talk about the big data topic, they generally do so following sensational cases, such as the Cambridge Analytica scandal of the past few months. The relationship between big data and the processing of personal data is not simple, because it is difficult to try to balance the value of technologies and the value of privacy. The biggest problem is related to data processing because, once inserted in the different systems, it is not always easy to understand by whom and how they are used, and with what data they can be integrated. The General data protection regulation, better known as GDPR, in application since last May 25, is changing the scenario at European level, but it is still early for an assessment of what will actually change. Unfortunately, every attempt to establish strict privacy regulations has a fundamental problem: the rules apply to well-defined geographical areas, each individual nation generally has different rules,

References

1. Bakshi, K. (n.d.). Technologies for Big Data. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 1-22. doi:10.4018/978-1-4666-4699-5.ch001
2. Sakr, S. (2018). NoSQL Database Systems. Encyclopedia of Big Data Technologies, 1-6. doi:10.1007/978-3-319-63962-8_50-2
3. Sakr, S. (2019). NoSQL Database Systems. Encyclopedia of Big Data Technologies, 1193-1198. doi:10.1007/978-3-319-77525-8_50
4. Kim, J. (n.d.). Big Data Sharing Among Academics. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 177-194. doi:10.4018/978-1-4666-4699-5.ch008
5. Tian, Y. (2018). Hybrid Systems Based on Traditional Database Extensions. Encyclopedia of Big Data Technologies, 1-6. doi:10.1007/978-3-319-63962-8_254-1
6. Tian, Y. (2019). Hybrid Systems Based on Traditional Database Extensions. Encyclopedia of Big Data Technologies, 986-992. doi:10.1007/978-3-319-77525-8_254
7. Shapiro, M., & Sutra, P. (2018). Database Consistency Models. Encyclopedia of Big Data Technologies, 1-11. doi:10.1007/978-3-319-63962-8_203-1
8. Loboz, C. (n.d.). Excess Entropy in Computer Systems. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 397-414. doi:10.4018/978-1-

4666-4699-5.ch016
9. Research and Implementation of Secure Storage of Big Data Based on Distributed Database. (2020). 2020 5th International Conference on Technologies in Manufacturing, Information and Computing (ICTMIC 2020). doi:10.25236/ictmic.2020.017
10. Analysis of Software Engineering Database Programming Technology in Big Data Age. (2020). 2020 5th International Conference on Technologies in Manufacturing, Information and Computing (ICTMIC 2020). doi:10.25236/ictmic.2020.058
11. Big Data Management, Technologies, and Applications. (2014). Advances in Data Mining and Database Management. doi:10.4018/978-1-4666-4699-5
12. Jung, G., & Mukherjee, T. (n.d.). Synchronizing Execution of Big Data in Distributed and Parallelized Environments. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 47-71. doi:10.4018/978-1-4666-4699-5.ch003
13. Naeem, M. A., Dobbie, G., & Weber, G. (n.d.). Big Data Management in the Context of Real-Time Data Warehousing. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 150-176. doi:10.4018/978-1-4666-4699-5.ch007
14. Tria, F. D., Lefons, E., & Tangorra, F. (n.d.. Big Data Warehouse Automatic Design Methodology. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 115-149. doi:10.4018/978-1-4666-4699-5.ch006
15. Yıldırım, A. A., Özdoğan, C., & Watson, D. (n.d.). Parallel Data Reduction Techniques for Big Datasets. Big Data Management, Technologies, and Applications Advances in Data Mining and Database Management, 72-93. doi:10.4018/978-1-4666-4699-5.ch004
16. Ameri, P. (2016). Database Techniques for Big Data. Big Data, 139-159. doi:10.1016/b978-0-12-805394-2.00006-4
17. R.R. Nadikattu. 2016 THE EMERGING ROLE OF ARTIFICIAL INTELLIGENCE IN MODERN SOCIETY. International Journal of Creative Research Thoughts. 4, 4 ,906-911.
18. R.R. Nadikattu. 2017. The Supremacy of Artificial intelligence and Neural Networks. International Journal of Creative Research Thoughts, Volume 5, Issue 1, 950-954.