Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 155 (2019) 599–605

Procedia Computer Science

www.elsevier.com/locate/procedia

International Workshop on Applying Data Mining Techniques to E-Learning and Pedagogical Approaches (ADMEPA)
August 19-21, 2019, Halifax, Canada

# Big Data Approach and its applications in Various Fields: Review

Loubna Rabhi[a] *, Noureddine Falih[b], Abdlekbir Afraites[a], Belaid Bouikhalene[b]

[a]*Mathematics and Applications Laboratory, Faculty of Sciences and Technology, Sultan Moulay Slimane University, Beni Mellal 23000, Morroco*

[b]*Laboratory of Innovation, Applied Mathematics and Information Technologies Polydisciplinary Faculty - Sultan Moulay Slimane University, Beni Mellal 23000, Morroco*

**Abstract**

"Big Data" refers to the extensive heave of data having varied and complex structure that can't be managed by traditional data handling methods and techniques.
Now a day's Big data analytics is most preferably used in enterprises, it plays a vital role in various fields. This paper presents an explanation of these applications such as Telecommunication, Business Process Management (BPM) and Human Resources Management (HRM).



*Keywords:* Big Data; Big Data Analytics; Big Data Applications.

## 1. Introduction

The growth of big data is the outcome of recent technology innovations. Several authors defined "Big Data" as huge amount of data coming from heterogeneous sources at a very high speed, that it is not possible for traditional tools and techniques to analyze and extract value from it.

Analytics of Big Data refer to the process of extracting insight from data raw by examining and analyzing its behavior and patterns using qualitative and quantitative techniques.

* Corresponding author. Tel.: +212-610-081-260; fax: +0-000-000-0000 .
*E-mail address:* rabhi.lubna@gmail.com

1877-0509 
10.1016/j.procs.2019.08.084

Recently, this large volume of big data cannot be analyzed by traditional tools, rather advanced Techniques called big data analytics. Sophisticated analytics can substantially improve decision making, minimize risks, and uncover valuable insights from the data. So, decisions need to be augmented by analyzing huge datasets using big data techniques and technologies. Big Data Analytics have gained great importance in the last years and are increasingly used in several contexts [1]: Healthcare, Public sector administration, Banking [2], Human Resources Management, Telecommunication...

This paper tries to cover major aspects of big data Analytics. It begins with a brief introduction of the topic. Important characteristics of Big Data is discussed in II. The issues of data Analytics is discussed in III. Transition from traditional data Analytics to big data Analytics is discussed in IV. And applications of Big Data in various sectors is treated in V, while the conclusion and future trends are drawn in VI.

## 2. Characteristics of Big Data

'Big Data' can be defined by the huge volume of data generated, in real time, from various digital sources like sensors, smartphones, social media and others. These data can have many different types as videos, audio, images, text and so on [3]. Several researchers are defined 'Big Data' by its Volume, Velocity and Variety (3Vs definition). That is means respectively, according to authors in [4], data size is large, the data will be created rapidly and the data will be collected in multiple types and captured from different sources:

- **Volume**: Denotes the large amount of data which is generating in every second from different sources.
- **Velocity**: Measures the speed at which data can be collected, analyzed and exploited.
- **Variety**: Means that the incoming data can have different types as: Structured, Semi Structured, Unstructured and Multi Structured.

Lakshen and al. told about 5Vs definition of big data instead of 3Vs. They added in [5] Value and Veracity:

- **Value**: The useful data among this large Volume of data
- **Veracity**: Denotes the correctness and accuracy of the data.

Also, according to research in [6], four others characteristics of Big data are broached to build 9Vs:

- **Variability**: It refers to data whose meaning is constantly changing.
- **Validity**: Means the data is correct and accurate for the intended use.
- **Volatility**: Means how long does company need to store data.
- **Visualization**: Presentation of data in readable and accessible manner for better decision making

To summarize, Big data are massive and rapidly-expanding, but it's also noisy, messy, constantly-changing, in hundreds of formats and virtually worthless without analysis and visualization.

## 3. Big Data Analytics

### *3.1. Big data lifecycle*

Big Data follows the lifecycle shown in the figure below [Fig. 1] from Data Generation to Data Visualization through Data Acquisition, Data Storage, and Data Analytics.

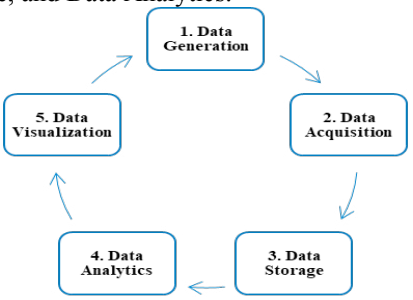


Fig. 1: Big Data life cycle

- **Data Generation**: Collect of data from various sources (sensors, video, click streams …) [7].
- **Data Acquisition**: The process of obtaining information from data by:
  (a) *Selection data*: Select pertinent data which is useful for the analysis
  (b) *Pre-processing data*: detect, clean, and filter the unnecessary and inconsistent data.
- **Data Storage**: persistently storing.
- **Data Analytics**: Analytics refers to the process of deriving actionable insights using qualitative and quantitative techniques [8] [Fig. 2].

(a) *Data Transformation*: After Gathering, Selection and pre-processing data, transforming preprocessed data into data-mining-capable format is required.

(b) *Data analysis*: After transforming data, analysis can be done using various statistical methods and data mining algorithms such as regression, classification, clustering … [9].

- Data Visualization: Representation of data insights in an interactive way going through:
  (c) *Evaluation*: Measure the results of data analysis;
  (d) *Interpretation*: displaying the output of data analysis by an interactive way.

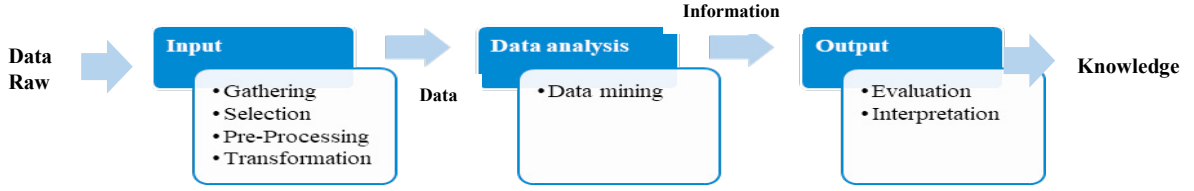


Fig. 2: Data analytics process

### *3.2. Big Data Analytics*

Nowadays, the data that need to be analyzed are big, contained heterogeneous data types, and even including streaming data which may change the statistical and data analysis approaches. Therefore, Traditional tools cannot be able to analyze this category of data. [10]. So, New approach of big data called 'Big Data Analytics' was born.

"Big data analytics" refers to advanced technologies designed to work with large volumes of heterogeneous data in order to improve the traditional Data Analytics Process mentioned in [Fig.2].

A lot of researchers talk about sophisticated types of analytics techniques as shown [Fig.3]: Descriptive Analytics, Predictive Analytics and Prescriptive Analytics:

- **Descriptive Analytics**: Gives information about What happened. In this technique, based on historical data, new insights are developed using statistical descriptions (such as Statistic Summary, Correlations and Sampling…) and Clustering (such as K-means…) [11].
- **Predictive Analytics**: Predicts the future outcomes using new statistical methods and predictive algorithms such as 'Decision Tree'. It provides information on what is likely happen in the future and what actions can be taken [12].
- **Prescriptive Analytics**: It is a type of predictive analytics. It helps to derive a best possible outcome by analyzing the possible outcomes by responding to the question So what? Now What?

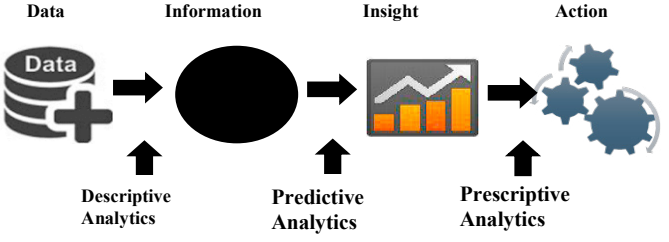


Fig. 3: Types of Analytics Techniques

## 4. Big Data Applications

### *4.1. Big Data and Business Process Management (BPM) Synergy*

Business Process (BP) is a succession of activities designed by human and systems that intends to achieve business goals [14]. Business Process Management (BPM) is a way to collect and treat data outcoming from processes in real time to support decision making. The lifecycle of BPM Project contains successive steps as it's shows in [Fig. 4].

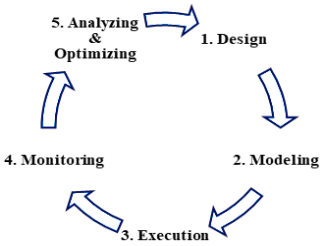


Fig. 4: BPM lifecycle

- **Design**: Modeling the design of what is currently being used and what will be used.
- **Modeling**: Analyzing process and performing "what if" analysis and then comparing the various process options to determine optimal improvements.
- **Execution**: Once the processes have been designed and simulated, they will be integrated into the information system for execution.
- **Monitoring**: Managing and supervising the processes.
- **Analysis and optimization**: Iterate for continuous improvement.

By the appearance of Big Data, organisations will need to integrate mobile data, social networks, digital video, and sensor data into their Business Process. So, they must be aware of Big Data challenges in order to make an efficient and intelligent BP that is aims to bring huge value for process decision makers and process actors. Some decisions are based on subjective judgment however, increasingly important decisions need to be based on hard data based on big data analytics. This is why big data analytics play a critical role in BPM by [13] [Fig. 5].

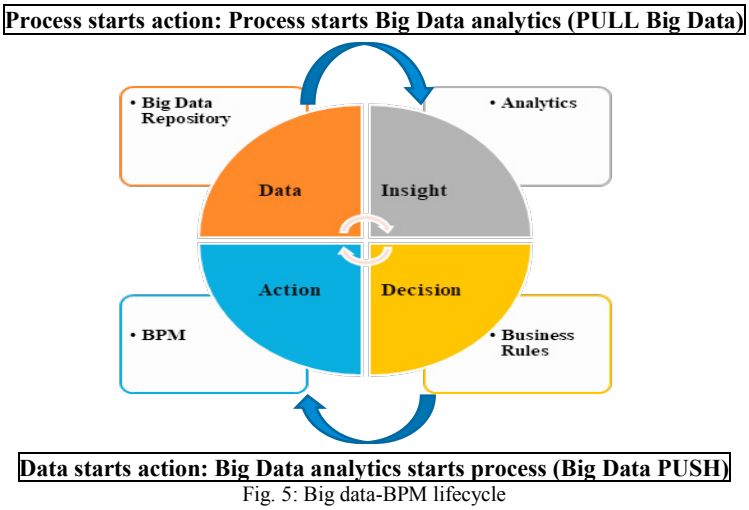


Fig. 5: Big data-BPM lifecycle

The big data opens up BPM to new data sources and opportunities to analyze processes more deeply and simulate potential improvements:

- **Big Data Repository**: refers to the large volume of data from different sources.
- **Analytics**: means a collection of reports, graphs and other analytical capabilities (Computation, Visualization...)

- **Business Rules**: After analytics, Managers can create "business rules" to make decisions that can be exercised through human judgment or fired automatically.
- **BPM**: Once all the decisions are made, the design and execution of the BP, in the opportune moment, can be triggered.
- **Learning from experience**: The massive amount of data transformed into actionable knowledge needs to be stored and analyzed in order to learn from experience and to continually change and improve the algorithms used to support decision making based on real metrics.

To conclude, Big Data Analytics / BPM synergy appears clearly in [14]:

- **More Insight:** Generally, more analysis should lead to more actionable insight;
- **Process Improvement:** Some processes can explicitly benefit from the results of Big Data analytics. There is no place for instinctive feeling decision while now decision can be supported with large-scale analysis;
- **Predictive Analytics:** To predict which inputs contribute to successful outputs or the opposite;
- **Learning from experience:** Business Process work may have already identified key metrics and these data can be used as input for Big Data implementation instead of reworking on global big data;
- **Process starts action (PULL Big Data)**: Process starts Big Data analytics;
- **Data starts action (Big Data PUSH**): Big Data analytics starts process.

### *4.2. Big Data and Human Resources Management (HRM) Synergy*

Big Data in Human Resources Management refer to the large volume of employees, customers, and transactional data in organizations. By analyzing these Big data by specific tools, HR processes (Recruitment Process, Training Process, Employees Career Management Process…) become more relevant which help companies to make decisions and create bigger benefits.

- **Recruitment Process:** Sometimes traditional recruitment process leads to bad hires Because most of the time, the interviewer may not have correct information about candidates that's leads to false results. Combining big data collected and traditional recruitment process may help recruiter to search for potential talents and everything he would possibly want to know about them is on their profile (personal picture, living conditions, social relationships…etc.) [15]. HR manager can match between the candidate's skills and personal beliefs and the company's needs. Hence, company can avoid invest in bad hires.
- **Training Process:** As it's known, talent training can lead to increase employees' level of knowledge and skill, it can also enhance their work performance. By traditional talent training organized by the company, professional trainers can be hired to ensure training which is spend a lot of material, human and financial resources. Usually such training takes traditional form of classroom instruction which cannot meet all the needs of employees. Using Big data context, any employee can easily search and access to the information that he needs to know on Internet at anytime and anywhere. [16]
- **Employees Career Management Process:** By analyzing all the gathered information of employees such as: interest on job, professional experience, performance …, HR Manager could find new ways to motivate employees and make them more engaged. Companies can combine traditional career management and career management of Big Data to make planning of new more effective talent-retention programs and to avoid employee turnover.

### *4.3. Big Data Approach for Telecommunication sectors*

#### A. Big Data opportunities in telecom sectors

- **Boost customer experience:** To increase customer loyalty, the operator needs to have an overview of its traceability: its web user account, its behavior towards his phone (web browsing, downloads...), his missed calls and its shares on social networks... [17]. Today, thanks to Big Data Approach, Telecom operators can collect this information from customers' mobiles. Operators can manage large subscriber databases that get richer each time a customer calls, creates a text message or uses the Internet...By using these databases, operators can:

a) *Predict churn*: Traditionally, Telecom operators could detect dissatisfaction of one of its subscribers if he has complained or if he has abandoned the service or thanks to purchase history. Today, with Big Data Analytics Techniques, operators can reveal new indicators of customers unsubscribing. They can identify those who are looking for new services on competitors' websites or posting negative contribution on a forum or on social networks...So, the operator can predict customers' churn by targeting "at risk" customers [18].

b) *Detect upselling opportunities*: Today, using Big Data Analytics Techniques, Telecom operators can get a global idea about purchase of their customers, they can propose to them a product or a service slightly higher and more expensive than their interests and make them aware of the ability to get even more of what he looking for.

c) *Identify cross-selling opportunities*: Through e-commerce websites, Telecom operators can collect 'Big' Data on customers purchase. They can then make aware them of ways to accessorize their purchase.

- **Network Optimization:** Today, with big data analytics techniques, operators seek to collect, store and analyze data generated by user devices and network devices like: routers, switches, base stations, and so one [19]. All of these data and others Key performance Indicators (KPIs) can help operators to well monitor network performance to ensure proper operation without problem.
- **Data Monetization:** Some of telecommunication companies see these massive volumes of data as a marketable resource to generate new revenue by aggregating and selling these data, they search to turn that data into money [20]. Many telecom operators also seek to commercially exploit this customer information, i.e. to generate new revenue by aggregating and selling this data. Some of them consider this an excellent financial opportunity.

To sum up, promoting loyalty, anticipating and reducing churn, offering upselling and cross-selling, optimizing network and personalization services are key areas where telecom operators can take advantage of Big Data Analytics.

### B. Big Data challenge in telecom sectors

Certainly, Big Data have opportunities in the telecommunications sectors. However, it also has divers' challenges such as: Big Data process, Real time Analytic, security and Revenue losses challenges [21].

- **Big Data process and Analytics challenge:** The complexity of the data to be managed today is a real challenge. Traditional tools of treatment of data seems simple compared to the contextual big data collected.
- **Security and data governance challenge:** The volume and diversity characteristics of Big Data lead to a new level of complexity in data security when telecom operators integrate new sources of information.
- **Revenue losses challenge:** Because of the competitive environment, gradual and continuous loss of landlines and increase of social networks, telecom operators have lost a large amount of revenue.

In fact, Big Data is still in the development and its related techniques and tools are far from mature. So, Human Resources Management, Telecom Sectors and others face also challenges in the use of Big Data in terms of Storage, Analytics and Management.

## 5. Conclusion and Future Trends

The authors of this paper have examined the innovative topic of "Big Data". They aim to reflect on big data analytics and its applications in various sectors such as Telecommunication, Business Process Management and Human Resources Management. This paper first defined what is meant by big data to consolidate the divergent speech in this topic. Second, it traits data analytics by discovering the data analytics process which is contains operations of selection, preprocessing, transformation, data mining, and interpretation/ evaluation. Third, it puts the accent on "Big data analytics" where Sophisticated analytics techniques are applied on big data in order to store, analyze and treat this extensive heave of enormous Data. Fourth, it applies Big Data Approach on various fields.

Although major innovations in analytical techniques for big data have not yet taken place. Our future work is planned in the sense of the contribution to the good governance of "Big Data Analytics Systems" since Big Data is still in the development and its related techniques and tools are far from mature. For instance, real-time analytics will likely become a profitable field of research because of the remarkable growth in location-aware social media and mobile apps.

## References


[1] S. Sagiroglu and D. Sinanc, "Big data: A review," 2013 Int. Conf. Collab. Technol. Syst., pp. 42–47, 2013.
[2] U. Srivastava and S. Gopalkrishnan, "Impact of big data analytics on banking sector: Learning for Indian Banks," Procedia Comput. Sci., vol. 50, pp. 643–652, 2015.
[3] M. A. Memon, S. Soomro, A. K. Jumani, and M. A. Kartio, "Big Data Analytics and Its Applications," vol. 1, no. 1, 2017.
[4] J. Campos, P. Sharma, U. G. Gabiria, E. Jantunen, and D. Baglee, "A Big Data Analytical Architecture for the Asset Management," Procedia CIRP, vol. 64, pp. 369–374, 2017.
[5] G. A. Lakshen, S. Vranes, and V. Janev, "Big data and quality: A literature review," 2016 24th Telecommun. Forum, pp. 1–4, 2016.
[6] S. S. Owais and N. S. Hussein, "Extract Five Categories CPIVW from the 9V's Characteristics of the Big Data," Int. J. Adv. Comput. Sci. Appl., vol. 7, no. 3, pp. 254–258, 2016.
[7] K. S. Jadon, R. S. Bhadoria, and G. S. Tomar, "A Review on Costing Issues in Big Data Analytics," Proc. - 2015 Int. Conf. Comput. Intell. Commun. Networks, CICN 2015, pp. 727–730, 2016.
[8] C. W. Tsai, C. F. Lai, H. C. Chao, and A. V. Vasilakos, "Big data analytics: a survey," J. Big Data, vol. 2, no. 1, pp. 1–32, 2015.
[9] M. Dave and H. Gianey, "Different clustering algorithms for Big Data analytics: A review," Proc. 5th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2016, pp. 328–333, 2017.
[10] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, "Next-Generation Big Data Analytics: State of the Art, Challenges, and Future Research Topics," IEEE Trans. Ind. Informatics, vol. 13, no. 4, pp. 1891–1899, 2017.
[11] G. Park, L. Chung, L. Khan, and S. Park, "A Modeling Framework for Business Process Reengineering Using Big Data Analytics and A Goal-Orientation," 2017.
[12] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," Int. J. Inf. Manage., vol. 35, no. 2, pp. 137–144, 2015.
[13] A. Hassani and S. Ayachi, "ScienceDirect ScienceDirect A framework for Business Process Data Management based on Big A framework for Business Data Process Data Management based on Big Approach Data Approach," Procedia Comput. Sci., vol. 121, pp. 740–747, 2017.
[14] M. Jha and L. O. Brien, "Combining big data analytics with business process using reengineering," IEEE Tenth Int. Conf. Res. Challenges Inf. Sci., 2016.
[15] J. Couaillier and D. U. Projet, "CATÉGORIE - ANALYTICS RH," pp. 50–54, 2016.
[16] S. Zang and M. Ye, "Human Resource Management in the Era of Big Data," J. Hum. Resour. Sustain. Stud., vol. 3, no. March, pp. 41–45, 2015.
[17] "How Big Data and Analytics can Drive Profitability for Telecom Operators — Exastax," exastax, 2017. [Online]. Available: https://www.exastax.com/big-data/how-big-data-analytics-drive-profitability-for-telecom/. [Accessed: 03-Mar-2018].
[18] Y. Huang et al., "Telco Churn Prediction with Big Data," Proc. 2015 ACM SIGMOD Int. Conf. Manag. Data - SIGMOD '15, pp. 607–618, 2015.
[19] C. M. Chen, "Use cases and challenges in telecom big data analytics," APSIPA Trans. Signal Inf. Process., vol. 5, no. 2016, pp. 1–7, 2016.
[20] J. Haapaniemi, "Jani Haapaniemi," no. October, 2017.
[21] B. Analytics, "Application du Big Data à l ' univers des télécommunications," 2013.