RESEARCHERID
THOMSON REUTERS

[Saidi* *et al.*, 6(9): September, 2017]
IC™ value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
Cod: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCE & RESEARCH TECHNOLOGY

WHY THE APPROACH BIG DATA, IS ESSENTIAL FOR THE RELIABILITY-CENTERED MAINTENANCE

Rabiae Saidi[*1] & Aziz Soulhi[2]

[*1]PhD Student at National Higher School of Mines, Laboratory LASTIMI, CEDOC EMI, Rabat, Morocco

[2]Professor at National Higher School of Mines, Rabat, Morocco

[2]ORCID: 0000-0003-1904-513X and the Scopus Author ID: 42162322200

DOI: 10.5281/zenodo.886903

ABSTRACT

This article aims to show why the approach Big Data is indispensable for the reliability-centered maintenance (RCM), and this as a result of volume as well as the diversity of data relating to the maintenance.

KEYWORDS: Big Data, RCM, Internet objects, Distributed Control System (DCS)

I. INTRODUCTION

The performance of an organization is linked to its ability to problem resolution and this to reduce see eliminate the gap between a situation of reference and a real situation. This is particularly valid for industrial equipment which will degrade as and and become faulty. However, thees the classical methods of resolution of problems adopted in the framework of the maintenance [1] are not sufficient, in view of the great mass of data derived from industrial equipment, the existence of several types and formats of these data from the instruments of measurement and control which issue the analog measures and a logical time allocated to the different parameters of the marches of the said equipment (flow, temperature, pressure,…) [2]. Therefore, a significant flow of data will be difficult to interpret in using those tools that will necessarily be reductionist in this case of the figure. Through this article, we intend to show why the big data is currently essential for or the reliability-centered maintenance (RCM).

II. LITERATURE REVIEW

1. Big Data

Big data as a concept

The concept of "Big Data" takes several forms through a set of definitions proposed mainly by large firms which operate on the sector of the informatics (information system, databases, business intelligence, mining). In this framework, it is wise to put the emphasis on the work of authors such as Jonathan Stuart Ward and Adam Barker [3] which have concluded that the big data refers to the storage and analysis of databases to both important and complex on the basis of a series of techniques such as NoSQL MapReduce, and **the automatic learning "machine learning"**. Here is a summary of those definitions that have been proposed to the concept "Big Data".

According to Gartner [4], the concept "Big Data" is strongly linked to three(3) aspects, called the famous "3V" to know, an important size of data(volume), the speed according to which these data are produced (**V**élocité) , as well as the diversity of formats of said data(**V**ariété), while the analysis of the data taking into account these three aspects must be carried out using the advanced tools capable of processing a **V**olume so this type of data characterized by these three "V".

As to Oracle [5], it guides the definition is Big Data A new management methodology which integrates traditional relational databases by other sources of data, both structured and unstructured (blogs, social networks, the networks of sensors, of image data, GPS data,..) and which are defined by the four "V" (volume, velocity,variety,value) These data

requiring an analysis of the data in depth. Has an indicative and non-limiting, this definition which is based necessarily on the technologies such as NoSQL, Hadoop .
For Intel [6], the concept "Big Data" is conditioned by the quantity of data available in the organizations producing a median of 300 terabytes of data per week by affirming that the forms the more present relating to data analyzed transactions are stored in relational databases, follow-up of documents, e-mail, the data from the sensor, blogs and social media.
While Microsoft [7]. The concept "Big Data" is used to express the process of the application of the power of calculation for the automatic learning on the basis of the massive data and complex.
To this we add other definitions; according to SAS [8], the Big Data Describes the large volume of data that are produced massively within a company. This volume of data represents a mine of investigation. In new techniques of analysis enabling by following the analysis and investigation of the opportunities for the benefit of the company.
According to IBM [9], the concept Big Data means the greatest amount of data produced most of the time in real time, and having a wide variety of sources such as the sensors, the networks of social media, … and which the storage capacity beyond that offered by the traditional relational databases requiring techniques of analysis advanced as the machine learning.

Big Data as a technology

Given that the traditional technologies (business intelligence, relational database, conventional servers,…) are exceeded in terms of performance , it is necessary to have technologies to optimize the time of processing and analysis of large databases , and therefore respond to the problem of "3V" characterizing the big data. Hear us quote mainly NoSQL and Hadoop [10],[11]:

- The NoSQL databases as MongoDB, Cassandra or Redis present storage solutions widely more efficient than the classical solution SQL for the analysis of heterogeneous data with important volumes and a speed of rapid increase tou in ensuring time to processing of queries reduced.
- The Hadoop technology is an infrastructure of servers for massively parallel processing (Distribute the treatments on multiple nodes) with a possibility to have the calculations made from the massive data divided by for example using the algorithm MapReduce[12].

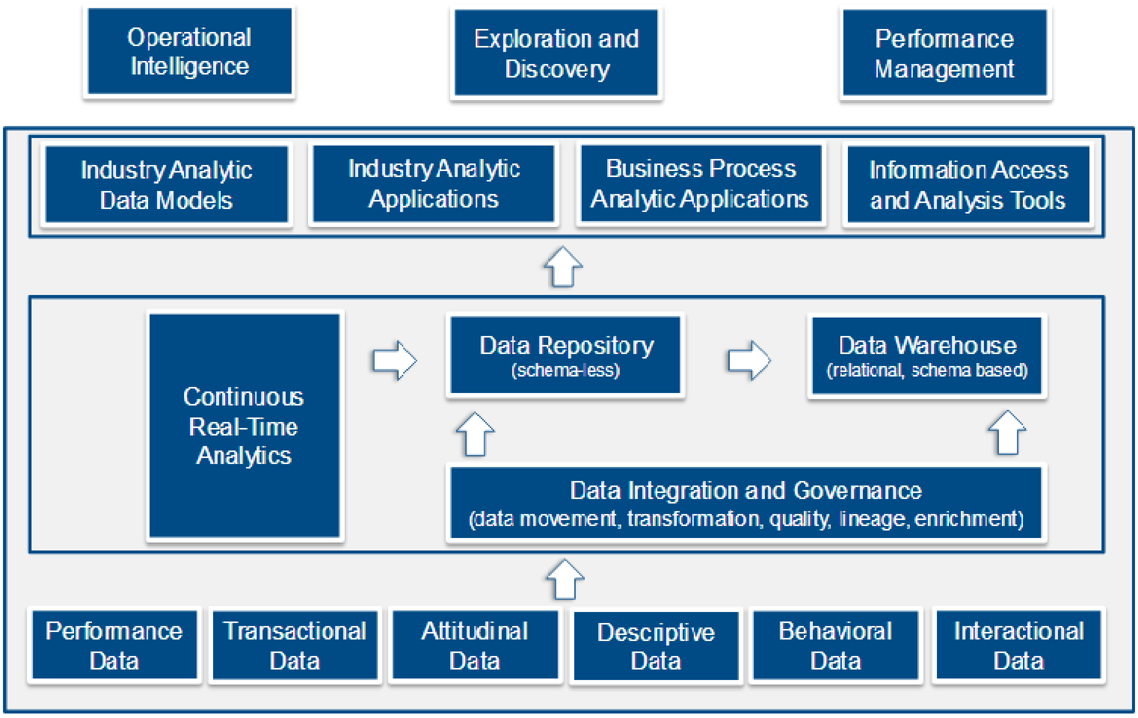

Figure 1 : Big Data Architecture [13]

Big Data as a methodology

Based on the literature, Big Data cannot be considered as a methodology, as there is no methodology cleanly there linked. However, it is appropriate to quote the méthodologie CRISP-dm (acronym of CRoss-Industry standard-

ISSN: 2277-9655
Impact Factor: 4.116
Cod: IJESS7
for-Data Mining) as being a specific methodology for the realisation of the projects Mining [14], and which is also adapted the Projects Big Data. It includes six (6) phases that begin necessarily the identification and characterization of the real problem and this before starting the other phases. In what follows, a brief description of the phases in question:

- **The characterization of the real problem** : it is the first phase which has for goal to characterize the real problem and well formulate the problematic of the operational business.
- **The collection of data** : Determine the data which are related to the real problem, taking into account the consistency between these data and their interpretation by report to the business.
- **The construction of the Data Hub** : It is a centralized database and structured which ensures the storage of all data carefully prepared for the analysis, on the basis of the initial data by making the classification of data according to specific criteria, their cleaning, the treatment of missing values in order to make them compatible with the algorithms that will be used. And avoid Especially that such algorithms give results in the case where the data are poorly prepared.
- **The Modeling** : C is the phase that requires concretely the competition and the competence of a data-scientistic (specialist in data science) in proceeding to test different appropriate algorithms and well adapted to the problem that should be addressed and this in the prospect of the good model. This phase typically includes three (3) sub-important phases, a first which is descriptive to generate the knowledge that reconstructs the past evolution of the process, a second which is predictive , explaining, as well as a third which is prescriptive to optimize a future situation.
- **The Evaluation** : This is the phase that allows you to ensure that the knowledge obtained through the model chosen meet the initial objectives which must be clearly formulated at the first phase (characterization of the real problem), and subsequently decide on the deployment of the model or, if necessary, to its improvement possibly in referring to the criteria determined (accuracy of the models obtained,…).
- **The deployment** : It is the relative phase to the production which allows to integrate the results of the modeling in the decision-making process and that the implementation in practice of the model chosen to be continuous.

Big Data as a tool of analysis

The analytical tools and conventional statistics are not adapted to analyze a large amount of data from different sources and to extract relevant information. By contrast, the Machine Learning proposes a set of modeling techniques based on the data and using advanced algorithms to exploit the large databases and perform predictions based on statistical techniques as predictive :

- The networks of neurons,
- The support vector machine,
- The classification techniques
- The decision tree
- Bayesian networks,….

2. Approach Reliability-Centered Maintenance

John Moubray [15] has developed the approach Reliability-Centered Maintenance II RCM II As an evolution of the first approach of the RCM method (**Reliability-Centered Maintenance**) which has been put in place at the beginning in the aeronautics sector and this for the improvement of the maintenance of the aircraft B747. It is an approach that has for objective the determination of what must be done to ensure that all equipment is able to accomplish what the user wants in a specific operational context and determined, in other words, determine with clarity and accuracy, what must achieve as a function for the said equipment is considered to be low within the meaning of the French Standard NF EN 13 306 which stipulates the definition of reliability. The RCMII approach is very pragmatic, she takes into consideration the main useful definitions for FMEA (analysis of the failure modes and their effects). Fundamental questions that should be asked for which it must answer to get the initial plan of maintenance. These questions relate to the functions, modes of operation and malfunction, the causes and consequences of failures and this to determine the maintenance tasks adequate, **as well as the importance of the consequences of the failures, the possibility to predict or prevent the dangerous failures with proactive maintenance and which can be implemented quickly in the case where no task of proactive maintenance may not be used**. In order to illustrate the conduct of a study of maintenance based on the reliability of type RCM2, the case of a palletising system involving chain conveyors and implanted in a plant for the manufacture of food cartons will be Treaty. The first part of the study presents the Table of FMEA specific to the RCM2 to identify the functions, functional failures, the modes and the consequences for the feeder to chains. Then, using the Logical Tree of decision of the RCM2, the maintenance tasks are presented for the six modes of failure of the first function and for the two failure modes associated with the second function. The intervals between maintenance tasks are also proposed. An analysis RCM II takes the form of Kaizen composed of people from the

ISSN: 2277-9655
Impact Factor: 4.116
Cod: IJESS7

operation, maintenance, management and technical resources such as engineering, health & safety and the environment. A marked benefits of team work is the rate of acceptance and belonging of management strategies of the deficiencies identified in the analysis. This allows a better understanding of the events leading to the failure of a equipment, the creation of a knowledge base, describing in detail the events being at the origin of the failures, the drafting a list of tasks applicable and effective potentially capable to eliminate the source of the events being at the origin of the failures as well as the extension of the life of expensive equipment. It is based on 10 key steps of an approach **:**

- Identification of the critical equipment.
- Choice of the analysis team
- Development of the operational context.
- Functional analysis of the equipment.
- Identification of functional failures.
- Analysis of the Failure Modes and Effects.
- Choice of the maintenance policy by mode of failure.
- Identification of intervals of search for failure.
- Identification of the periodicity of tasks and choice of resources.
- Presentation and implementation of the plan of actions.

3. Justification The usefulness of Big Data for Maintenance

The challenge of the maintenance in the face to the volume of data generated

One of the major challenges of the maintenance is the explosion in the volume of data. This is confirmed by example by Pierre MARCHADIER, vice-president of Dassault Systems by saying "…..*An engine on a Paris /New York generates up to 500 GB of data in flight, and the A350 when he flies today to make its tests is roughly 2 TB of data by flight* ." Therefore, the analysis of the data in relation to the maintenance generates a Triple problematic (**problem of 3V**)

- A large **V**olume of maintenance data.
- A **V**ariété of information (several sources unstructured and structured with different formats)
- A **V**élocité required in terms of the frequency of processing and analysis of data.

Link between the automatism of equipment and informatics

Most of the industrial equipment are equipped with sensors to measure and control enabling the acquisition of data in real time through a digital communication between computer networks and local networks of automation in accordance with the CIM model [16] (computer integrated manufacturing), which allows a segmentation of the automatism in several levels:

- Level 0 : Sensor/actuator,
- Level 1 : automatism of equipment,
- Level 2: Supervision at the level of a control room .
- Levels3/4: the exchange with the computer network

The level sensor/actuator requires a transfer in real time (a few milliseconds), and concerning the Continuous measures and binary characterizing the parameters of steps of a production system from the Field, whereas the level "factory" allows you to transfer data in large production to networks It. Each level of the CIM model, reflects the performance of the exchange / quantities of information either In real time, either in time deferred, so control of the machines is ensure by the networks of the levels 0.1, 2. The levels 4 provide rather the piloting of the process and the connection with the computer networks.

Industrial Internet of Things (IIoT)

The Internet of Objects refers to a network of physical objects which can be of the production equipment as well as the different measuring sensors and control integrated with electronics, software, sensors, and connectivity of the network, which allows these objects to collect and share massive amounts of data. With the advent of the Internet of objects, the different equipment of a production system can now communicate using industry standard communications protocols of the Internet, and therefore, the data may be available to be used to be used in the framework of the Artificial Intelligence, beyond the simple control of the process.

Figure 2 : Example Industrial interface The Internet Of thing (IIoT) [17]

Acquisition and Archiving of a large mass of data by the DCS (Distributed Control System)

In industry (as manufacturing the automotive sector and continues as the cement industry) All installations and industrial processes are related are supervised by the means of a DCS (Distributed Control System). This type of system has a human-machine interface through a network of digital communication (local network industrial) and which is equipped with a network architecture that can be specific to each publisher of these systems, a set of measuring and control instruments allowing to make in real time the acquisition of the parameters of the marches of facilities such as the flow, pressure, temperature. The history of all parameters of walking on a factory are available for a long period of up to six months depending on the editor of these DCS and this to have a mainly traceability of the analog inputs related to ongoing settings, logical entries for events as well as the logical outputs and analog for the control of the equipment and the instructions of the control loops, this clearly means that at the end of a certain period. It is certain to have a massive acquisition of data from the various sensors and production equipment as well as the need to have a large storage capacity of the historic.

Figure 3: Example of a control room [18]

Figure 4: Example of a DCS [19]

ISSN: 2277-9655
Impact Factor: 4.116
Cod: IJESS7

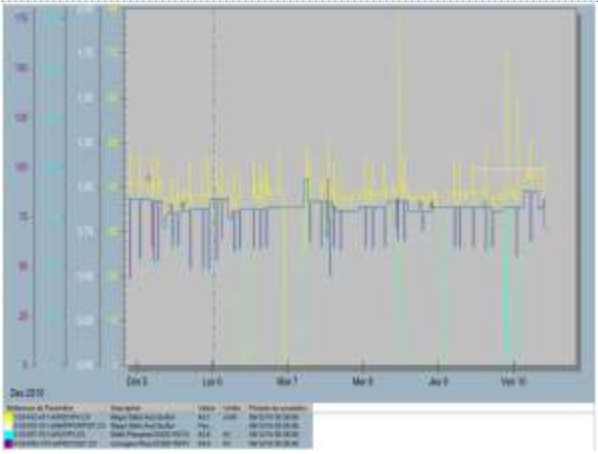

Figure 5 : Example of history from a DCS

Massive flows of acquisition of production data

This flow of data acquisition of production is reinforced by the acceleration of the automatic acquisition of data relating to production systems through the networks of land [20] which ensure the communication between the production equipment, the sensors as well as the control room in which, we find the stations of supervision. In this context, there is necessarily a exchange of information of different formats and sometimes in real time according to the type of the local network industry that is therefore built at different levels within the pyramid CIM. Therefore, **the**hen these sensors are connected to a communications network, this greatly increases the volume of data available for the analysis, and also has the potential to allow for an analysis in real time.

Acquisition of data by the technologies of predictive control of equipment

The tools of predictive control are used in the framework of the predictive maintenance in order to anticipate the detection of anomalies of production equipment and this before they actually produce of faults causing of forced outages. The state of the art in the subject fit that there three (3) main techniques used

- **The Acoustic Measurement** : this technique use of ultrasonic waves with a frequency exceeding 20Khtz, in order to rule on the status of abnormal operation of the equipment by detecting mechanical defects, sealing as well of leaks of oils,…
- **The thermography:** thermography allows to detect subject of temperature sensors mainly, infrared cameras used to Distance, the thermal variations of abnormal equipment as well as the electrical installations (low- voltage,…).
- **The Vibration Analysis**: This technique allows you to measure and analyze the vibrations emitted by a rotating equipment to detect defects of tightening, misalignments of the trees of transmission as well as the degradation.

Different formats of the data relating to the maintenance

The three (3) techniques already cited allow an acquisition of data semi-automatic, in other words, in a first time, the measurement devices automatically record the different measures carried out, then a manual storage is done other system of external information combining other data formats. Hear us quote essentially the tool production management assisted by computers(POAG) or even systems of technical data management(SGDT) [21] which also generates a significant stream of digital data (indicators of maintenance such as the MTBF, MTTR,..) as well as the textual data such as requests for intervention, the intervention reports. On the other hand, most of the production equipment are currently equipped with a set of instruments to measure and control the issuing of real time measures concerning their functions (power consumption, levels of oils…). This represents another source of automatic data acquisition.

ISSN: 2277-9655
Impact Factor: 4.116
Cod: IJESS7

Figure 6: Example CMMS Architecture

III. RESULTS AND DISCUSSION

In the prospect of having, a methodology suited to the implementation of a project "Big Data maintenance ". We can say that there are opportunities to combine the first two phases concerning the methodology crisp-DM, namely the phase of characterization of the problem and the one relating to the collection of data with the Kaizen approach aims to continuous improvement by the internal means a business [23]. This approach requires at the beginning a real involvement of employees who have mastered actually the field (phase Gemba Gembutsu of the Kaizen approach), and the industrial production system, the operation of the machines as well as their modes of Failures This involvement is very beneficial for the methodology CRISP-DM if we want to apply it to the case of the maintenance in order to identify with precision the problematic of the maintenance. On the other hand, when the collection of initial data, the employees of the field know all the textual sources as the CMMS, reports of intervention, the historical subject Excel files, as well as any of structured and unstructured data sources and which will be processed. This can be a contribution for the data-scientistic during the phases «Data Collection" and "Construction of the Data Hub".

With regard to the phase modeling, the data-scientist, may decide to use for example, the techniques of Machine Learning, however, this approach is based heavily on the knowledge of the experts. It is timely to formalise in the framework of a shipyard in KAIZEN bringing together such experts in the place where it would be necessary to make the analyzes to study the decision-making process in a participatory manner and this using a share of analytical tools such as fuzzy logic or Bayesian networks to make the prognosis and the requirements. which fits perfectly in the perspective of the maintenance based on the reliability (RCM), the purpose of which is to ensure the reliability at lower cost of which the major challenge is to develop plans for a sharp maintenance that could be optimized by a set of prescription emanating from sub-phases Description and prediction of failures of equipment.

Taking into account the foregoing, the combination of approaches KAIEZN and crisp-DM allows you to connect the human intelligence and collective (teamwork, knowledge of the process,) with artificial intelligence involving of predictions and classifications using the machine learning.

IV. CONCLUSION

Through this article, we have made the arguments stating that the concept of Big Data is indispensable in the approach “reliability-centered maintenance (RCM)”, with particular emphasis on the complexity and the diversity of data (Variety) and referring to the multiplication of sources. Most of the industrial equipment are equipped with a set of sensors delivering real-time automatic data of different formats; on the other hand, the CMMS systems produce a lot of data which are also well textual and digital. We quote, the reports of intervention are the entry is made through the CMMS, the history of Faults as well as the different indicators of the maintenance (MTBF, MTTR…).

On the other hand, we have mentioned that the sensors -which are mounted on most of the equipment modern industry- are connected to a communications network, this increases the volume of data available in real time. The deepening of this subject is now part of'd a doctoral thesis, which has as a perspective contribution in the implementation of the synergy between the Machine Learning and Maintenance.

ISSN: 2277-9655
Impact Factor: 4.116
Cod: IJESS7

V. REFERENCES

[1] J. Riout, Le guide de l'AMDEC machine, publisher CETIM, page 1999
[2] E. Gaudreau, B. Agard, Mr. Trepanier, P. Baptiste, "reactive Pilotage of production systems with the help of Intelligent sensors", 6th International Congress of Industrial Engineering, June 2015, page 4.
[3] Jonathan Stuart Ward and Adam Barker, "Undefined By Data: A Survey of Big Data Definitions", School Of Computer Science, University of St Andrews, UK
[4] www.gartner.com Consulted August 28, 2017
[5] www.oracle.com Consulted August 26, 2017
[6] www.intel.com Consulted August 28, 2017
[7] www.micrsoft.com Consulted August 28, 2017
[8] www.sas.com Consulted August 26, 2017
[9] www.ibm.com Consulted August 26, 2017
[10] C.L. Philip Chen, and C.-Y. Zhang, "data-intensive applications, challenges, techniques and technologies: A survey on Big Data", Information Sciences, 2014, 314-347
[11] www.lebigdata.fr Consulted August 27, 2017
[12] J. Lin , and C. Dyer,"Data-Intensive text processing with MapReduce", Graeme Hirst (Series Editor),2010
[13] D. Vesset, "The Importance of the Information Architecture, Sponsored by: Oracle", June 2015, page 6.
[14] c. Oprean, " State of the art on the methodological aspects and processes in Knowledge Discovery in Databases ", TELECOM Bretagne(France), January 27, 2010, page 7.
[15] J. Monbray, "Reliability Centered Maintenance", British Library, 1997, page 323.
[16] T. Manokaram,"Computer Integrated Manufacturing", Sasurie College of Engineering, 2015
[17] www.sap.com consulted the 21/08/17
[18] www.rockwellautomation.com consulted the 21/08/17
[19] www.electricaltechnology.org consulted the 21/08/17
[20] I Thomesse, networks of ground, LORIA-Institut National Polytechnique de Lorraine, acts of the Ecole Of Summer Real Time, 2005.
[21] A. Courtois, C. Bonnefous, and Mr. Pillet, "Management of the production "editions of Organization, 2003, 404-420 page

CITE AN ARTICLE

Saidi, R., & Soulhi, A. (2017). WHY THE APPROACH BIG DATA, IS ESSENTIAL FOR THE RELIABILITY-CENTERED MAINTENANCE. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY*, *6*(9), 217-224. Retrieved September 5, 2017.