Harnessing the power of data and event data for Business Process Improvement

Abir Ismaili-Alaoui, Karim Baïna, Khalid Benali

Harnessing the power of data and event data for Business Process Improvement

Abir Ismaili-Alaoui [a,b *], karim Baïna [a], khalid Benali[b]

[a]Alqualsadi, Rabat IT Center, ENSIAS, Mohammed V University,
Rabat, Morocco
[b]Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France

Abstract

Faced with a competitive and a continuous changing environment, traditional approaches that treat a company as a closed environment are no longer appropriate. To overcome this problem of isolation and non-communication, organizations tend to increasingly use Business Process Management - BPM. Recently with the rise of new technologies such as of big data, Internet of things, Cloud computing, etc, organizations are faced with many factors and challenges that generate real changes in the traditional BPM. Among these challenges, we have the huge amount of data and event data that are continuously gathered. Such data must be adequately exploited to extract high added value that can assist the organization in its decision making process. However, traditional BPM systems present different limits, as they do not facilitate the use of knowledge extracted from this data by business processes, because they do not benefit from statistical functionalities and data analysis and manipulation techniques. Several researches have been done in this area to link event data and data analysis to BPM by using, for example, process mining or machine learning algorithms. This paper shows how data and event data are the key to get a better understanding of the functioning of business processes, and to start the journey of business process improvement towards a stateful, context-aware, and proactive business process.

Keywords: Business Process Management, Data, Event data, Data Analysis, Machine Learning, Business processes Improvement.

1. Introduction

During its life cycle, an organization often goes through several stages starting with the foundation or birth, through growth or development to reach the stage of maturity and finish by either a decline or renewal. To avoid decline and degradation of service, the company must have a transversal vision of its various activities, an overall steering and a better understanding of its objectives; moreover the performance level of each organization is indirectly linked to the performance and the quality of their business processes. In this regard, managing effectively the business processes in any organization, can have a positive impact on the effectiveness and proper functioning of its activities, as it allows the company to have a clear vision of its objectives. However, companies nowadays are finding that the traditional BPM systems are no longer meeting the requirements of the market because it has different limits on the processing and analysis of large and diverse data and event data, in almost real time, in order to extract useful insights from it. Nowadays, data is considered by companies as a catalyst for innovation, development and continuous improvement of their performance, their services and their customer relationships.

The new digitalized era and the rise of several new technologies such as big data, fast data, cloud computing, Internet of Things (IoT), etc, implies new business process problems and challenges linked basically to the tremendous amount of data and event data that are constantly collected within the organization. These data represent for enterprises a real engine of growth. However, a large amount of raw data is not valuable; data must go through a whole process to extract value from it. The analysis of huge data helps organizations to extract information and then knowledge, because the real value is in how organizations will use that data and turn their organization into an information-centric company that relies on insights

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

derived from data analyses for their decision-making" [1]. Hence integrating the science of data in each step of business process life cycle is very crucial for the process improvement.

The remainder of this paper is structured as follows: important concepts in our work are presented in the next section. In section 3 we highlight the importance of data and event data for process improvement and we show how data science can be used as tool to achieve business process improvement. Section 4 is devoted to the presentation of our case study and experimental results. Section 5 concludes the paper.

2. Concepts and Definitions

2.1. Business process management BPM

M. Dumas, et al. [2] define BPM as "the art and science of overseeing how work is performed in an organization to ensure consistent outcomes and to take advantage of improvement opportunities". For Gartner [3]: "BPM is the discipline of managing processes (rather than tasks) as the means for improving business performance outcomes and operational agility". Whereas IBM [4] finds that: "BPM is a discipline that leverages software and services to provide total visibility into an organization. Discover, document, automate, and continuously improve business processes to increase efficiency and reduce costs". P. Harmon [5] in another side mentions that: "BPM is a management discipline focused on improving corporate performance by managing a company's business processes". Each one of these definitions focused on one aspect of BPM, whereas business processes management is a systemic approach involving the various stakeholders of the company, from strategic to technical level.

C. Moller et al. [6] focuses on all sides of BPM, by saying that "BPM is a holistic management discipline that uses technology to control and operate the entire business through rules that clearly defines business process. BPM is about continuous improvement and optimizing process to ensure high performance and by achieving agility and flexibility as a tool to gain competitive advantages". This definition shows that BPM is an open and flexible discipline that can interact with several technologies in order to achieve the goals predefined by each organization, and ensure an effective and proper functioning of its activities and to meet the requirements of the market [22].

2.2. Business Process Improvement

Business Process Improvement (BPI) represents any positive change that can be done at the level of business process and business process model. BPI is a systematic approach that enables organizations, once adopted, to continuously enhance their existing Business Processes. It is considered as "an intrinsic aspect of effective Business Process Management (BPM)" [7]. The Business Process Improvement was first used by H.James Harrington in 1991 [8] where he provides a guide for organizations to start the journey of process improvement based on the quality improvement techniques used basically in the manufacturing sector. He describes BPI in 1997 [9] as "a methodology that is designed to bring about step-function improvements in administrative and support processes using approaches such as process benchmarking, process redesign and process reengineering". Parallel to the concept of BPI, we have Business Process Re-engineering (BPR) that was introduced by Hammer in 1990 [10], it is considered as a radical methodology to redesign business processes from scratch, unlike BPI which aims to enhance the quality and the productivity of the existing ones.

According to S.Page in [11] "BPI does affect the entire business system, including the employees who do the work; the information technology systems that support the process; the measurements established to assess the effectiveness, efficiency, and adaptability of the process; and reward and recognition programs that exist in a company". Therefore, adopting this methodology within the organization is becoming mandatory in order to evolve its performance and achieve its business goals. Concretely, improving business process means improving the intrinsic characteristics of this process. Each business process is characterized by some specific performance metrics; A. Shtub et al. [12] introduced what they call the four dimensions of competitiveness: flexibility, quality, cost and time, those four cornerstones of competitiveness are leading to the survival of an organization and to its success. In the same perspective, those four main dimensions are widely used in BPR and BPI literature [13] [14] and known as the Devil's quadrangle; this concept was introduced in 1995 [15] by Brand and Van der kolk. In [16] SW. Cranenbroek explains each dimension as follow: The quality, which is divided in two types (internal quality and external quality). Internal quality refers to the social and psychological factors related to work. The external quality consists of the degree of how the products meet the requirements of the customers. Time, which can be represented in three main categories of time dimension according to Brand and van der Kolk : service time, queue time and

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

wait time. We have also flexibility that represent the degree organizations can adapt to changes in the environment. And finally, cost that can change based on the process type and resources (Human and machine) used in this process.

The purpose of Business Process Improvement according to [17] is to improve those metrics by enhancing the quality of a business process, reducing its response time and its cost. Whereas Gordon Masiuk in [18] adds 3 others metrics to those used by [17], he introduced 6 key performance indicators for processes: Quality, Cost, Time, Quantity, Satisfaction and value. G.Masiuk classifies the 6 keys on 4 categories of process measures: first we have the input measures which represent the assessment of products, services or information that feed into a process (quality, on time delivery). Second we have process efficiency measures which mean the assessment of how well the process is functioning (cycle time, task time, accuracy, completeness, cost). in third, there is the output result measures which represent the assessment of the outputs of the process (quality, meeting employee/team/stakeholder requirements, quantity/volume produced). And finally the outcome measures which means the assessment of the outcome (impact) of outputs on the business and stakeholders (stakeholder satisfaction, business value, productivity).

Improving business processes is no longer limited only to the quality of process models, or the business process performance (by improving the predefined metrics and measerements). Because nowadays, organizations are more intrested in achieving other aspects in their business processes such as adaptability, intelligence, flexibility and proactivity. In the literature we found several approaches that aim to improve the intelligent and proactivity aspects of business process based on different techniques such as: ubiquitous computing [17], knowledge management [16], Decision-making [17], E-Collaboration[20], Machine learning [21] [22] , Process mining [23], etc. Improving business process is a very open issue that still attracts more attention from the research community, most of these contributions focus on adding new Improvement techniques in order to manage correctly and successfully their business processes, and to make them more adaptable, proactive and smart. However , all these techniques have data as the important ingredient to achieve process improvement, that is why harnessing the power of these data can have a huge impact on business process improvement.

3. The science of data as a tool for process improvement

Data have a plethora of facets that can be exploited to help organizations enhance their performance. In a survey based on 325 responses [24], 70% of organizations consider (Big) Data as “an opportunity to discover new facts about their customers, markets, partners, costs and operations, and then they use that information for business advantage”. But Data is a double edge sword; it can either improve the quality of business process or deteriorate it, based on its quality and how we use it. In fact, there is a proportional relationship between the quality of data and the quality of process and subsequently the quality of decisions taken by the organization. This idea has been detailed and clarified by D. Opolon et al. in [25] as illustrates in Figure 2.

To take full advantage from Data, organizations must pay attention to the quality of data that they use. Because, accurate and consistent data merged to an improved business processes, can lead to an appropriate decision making. That is why data science is needed to achieve process enhancement in every domain. However real world data sets and event logs come most of the time in complex format, which make it difficult to tackle the research questions (predictive modeling, process analysis, process dicovery, metrics adjustments, …). In fact, tidy and preprocessed data are the starting points towards finding reliable models that can be used to describe these gathered data [28]. Raw data and event data exploration and analyses must go through “transform-visualize-model” loop as proposed by Wickham & Grolemund in [26] (see Figure 1). At the end of this loop, we have more accurate data that represent the cornerstone of 1) assumptions making 2) insights and predictions 3) and finaly actions and decisions making.

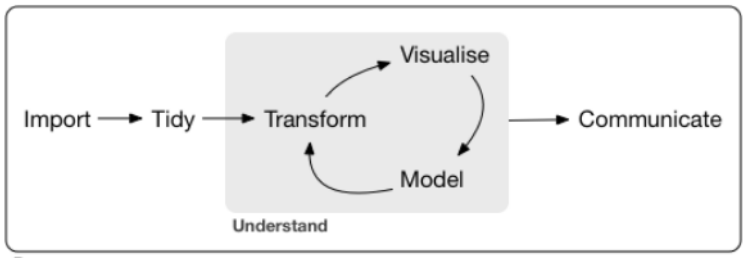

Figure 1: transform-visualize-model” loop

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

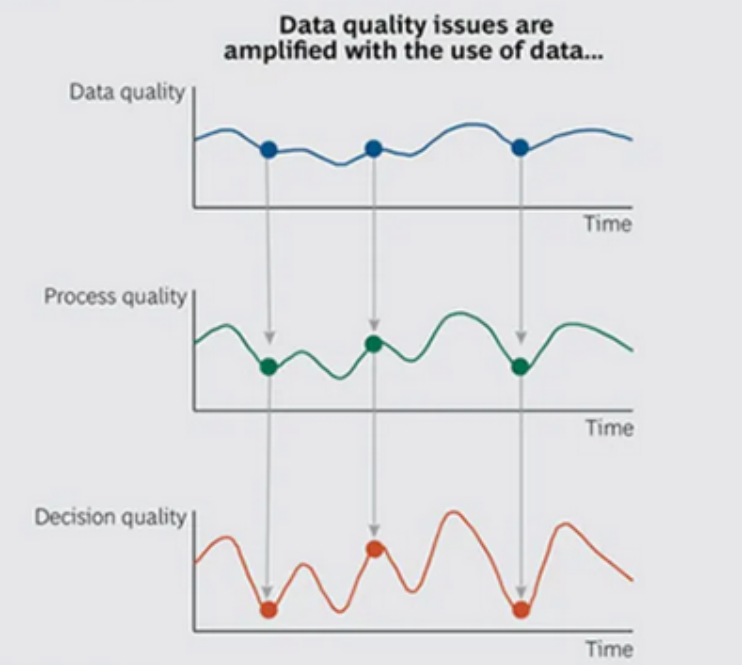

Figure 2: Impact of data quality on process and decisions

In the field of business process management, data that are gathered after the execution of each business process instance is called event data. The term event Data refers to actions performed by entities. Each event data must have at least three key pieces of information: action, timestamp, and state [27]. But for more accuracy, an event data must have at least 6 different pieces of information [28] (see Figure 3):

- Case identifier : the identifier of the case to which the event belongs. Where a case represents an instance of the process that we are dealing with.
- Activity identifier : the identifier of the activity that the event refers to.
- Activity instance identifier: the identifier of the activity instance.
- Transactional life cycle stage: the status in the transactional life cycle of the activity instance.
- Resource identifier: the identifier of the resource that execute the activities.
- Timestamp: the timestamp of the event.

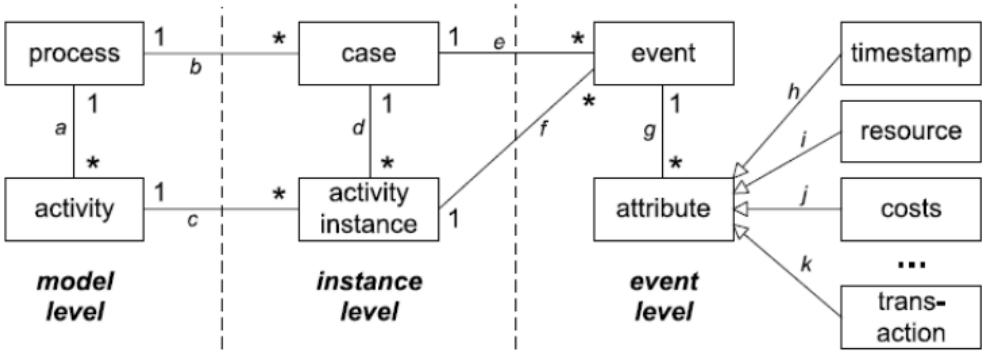

Figure 3: Event data model [30]

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

Nowadays, process mining is widely used in the field of BPM. It represents a set of techniques that are used to extract knowledge from event logs that are available in information systems, in order to discover, monitor and improve real processes [29]. However, the lack of tidiness in the gathered event data and its complexity can hinder the quality of process analyses, either through process mining or through other techniques. Hence, it is more interesting to combine some data science techniques to have tidy and accurate data and also to produce descriptive and exploratory insights to empower the business process improvement journey.

4. Case Study and Application

To confirm the efficiency of our assumptions we made a series of experiments based on a case study that belongs to silver economy domain [31].

Supervision and incident management business processes in health-care are considered as an event-driven business processes. The instances of these processes are, in some cases, launched by IoT generated events, especially in the current research fields related to health-care. Let us consider a video surveillance company that edits an automatic falls detection system for elderly people and offers a 24/7 automatic alert solution and a quick rescue without the intervention of the person in danger. The incident management process used in this case study is based on a real-time analysis of alerts received from 24/7 streaming cameras for detecting falls of elderly people. The events that launch our business process instances in this case study are generated by different cameras placed in each patient room in order to detect possible falls. During the launched instance of this business process, the incident (that triggers this instance) is qualified by human agents and classified during the qualification step into four categories (see Fig. 3): false alerts (empty place), false alerts (active person), true alerts with average risk level (seated person), and true alerts with high risk level (lying down person). The human resource determines whether an assistance action is necessary or not, that's why each received event require an in-time and vigilant qualification in order to prevent delayed intervention or incorrect qualification. Once the incident has been confirmed, a handling step will take place to take care of the patient that triggers this instance. After that the whole activities go through quality assessment step before closing the event.

This business process generates a huge amount a event data. Exploring these data in oder to take full advantage from it, for improving the business process, presents a real challenge for the organization.

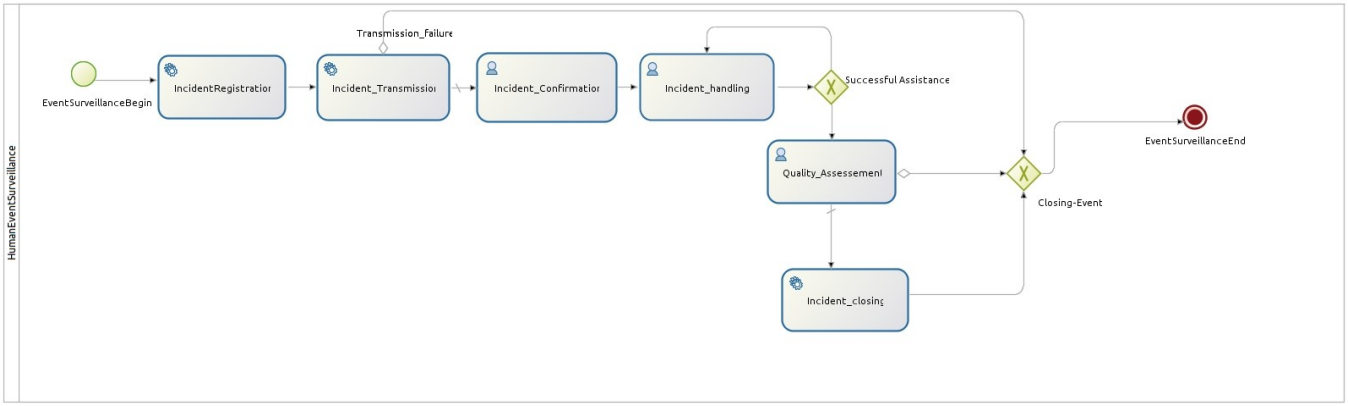

Figure 3: Process model of incident qualification and assessment

The event data gathered from our incident management process past instances conceal a huge amount of information and insights. And this data represent the main ingredient that we can use in each step of the process improvement journey. However, this gathered data come most of the time in shapes that makes it difficult to easily extract useful insights for the decision-making process. That's why it is mandatory to have recourse to data science techniques in order to harness the power and the abundance of these data.

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

The starting point of this analysis steps was the gathering of the generated data from our business process execution. This raw data (in csv files) must go through two important steps :

1 - Data tidiness, exploration and understanding.

2 - Transforming the data set of event data into a useful event log.

The historical data in this dataset of patients falls are gathered from our incident management process past instances (from 02-01-2017 to 12-06-2018), this dataset contains 5442 observations generated by several patients. To transform this raw data into a useful input for either statistics and machine learning algorithms or process mining algorithms, we have recourse to the open source statistical language R which is widely used by data science researchers. R is considered as a programming language and as a data analysis environment. The fast-growing and open community and the R-package ecosystem on CRAN (Comprehensive R Archive Network) have contribute to its recent impressive growth.

In our experiments we have used some interesting R packages for such as:

- Tidyverse : which is a coherent system of packages for data manipulation, exploration and visualization. It facilitates the data analysis workflow [32].

- BupaR: which is a collection of packages used for business process analysis, in order to explore and visualize event data, to create event logs from business process event data, and to monitor business processes [33].

The following table (see table 1) describes the event log that we have created from our initial csv file that contains the raw data from our business process execution.

Table 1: example of incident process event log

Event_Source	Patient	Timestamp	Assessment Steps	Assessment Steps_id	Status	resource
9Chambre112	8	2017-01-05 04:56:11	IncidentRegistration	1	start	HR1
9Chambre112	8	2017-01-05 04:56:15	IncidentRegistration	1	Complete	HR1
9Chambre112	8	2017-01-05 04:57:05	Incident_transmission	2	Complete	HR1
18Chambre106	110	2017-07-26 02:51:04	Incident_handling	4	start	HR4
18Chambre106	110	2017-07-26 02:55:14	Incident_handling	4	Complete	HR4
18Chambre108	20	2018-04-25 23:04:09	Incident_closing	6	Complete	HR7
6Chambre204	103	2018-04-30 09:59:14	Incident_confirmation	3	Complete	HR3

The tidy data set that we obtain from the first step and the event log that we created in the second step, make our data and event data more understandable and ready to use in order to extract useful insights and information, regardless of the intended objective of the research (process monitoring, scheduling, process mining, …). For example, by visualizing our business process we can have a lots of information about the control-flow perspective (control-flow deviation for example) of our process (see Figure 4) or from activities perspective (see Figure 5), and more other perspectives (resource, performance, processing time, throughput time, trace length, traces with self-loop, …).

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

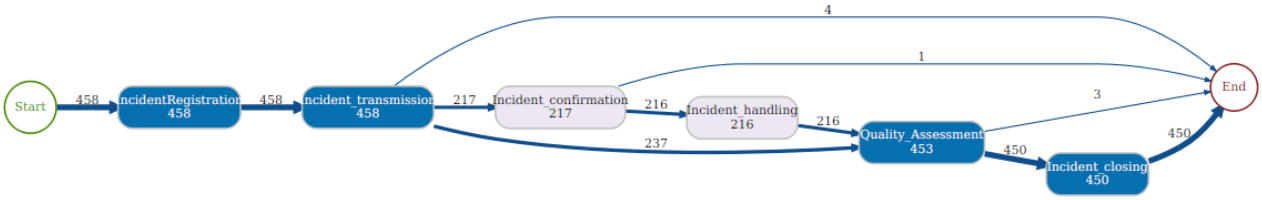

Figure 4: Incident qualification and assessment process map

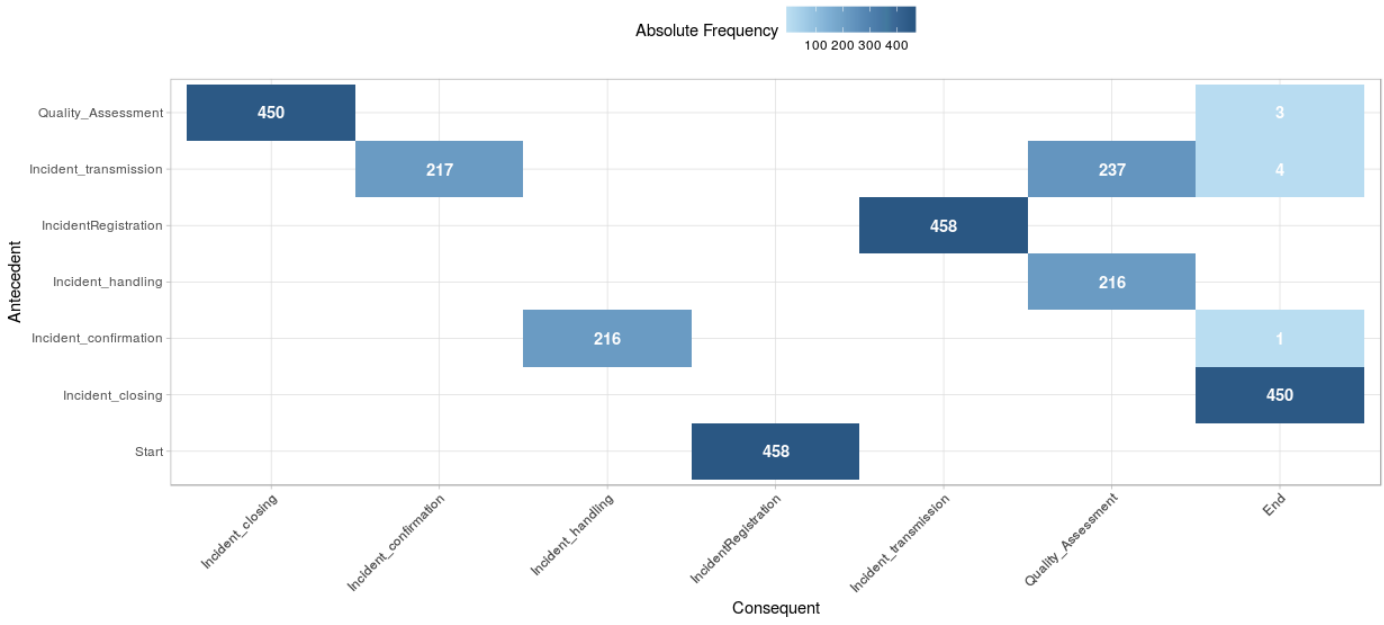

Figure 5: Incident qualification and assessment activities frequency

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

5. Conclusion

This paper shows how data and event data are the key to get a better understanding of the functioning of business processes. In fact, preprocessing and exploring data before using it help to get correct assumptions and insights in order to make correct predictions and finally take correct and accurate actions and decisions.

The motivation of this paper is to emphasize the importance of data and event data for any organization in order to manage correctly and successfully their business processes and start the business process improvement journey towards a stateful, context-aware, and proactive business process, in a solid and correct way.

REFERENCES

[1] M. Rijmenam, “Why the 3V’s Are Not Sufficient To Describe Big Data,” (DATAFLOQ), [Online], https://datafloq.com/read/3vs-sufficient-describe-big-data/166 (Accessed : june 2015).

[2] M. Dumas, M. La Rosa, J. Mendling and H. Reijer, “Fundamentals of business process management,” (pp. I-XXVII). Berlin: Springer, 2013.

[3] Gartner, [Online] 2013, http://www.gartner.com/it-glossary/business-process-management-bpm/ (Accessed: April-2015).

[4] “What is BPM?”, IBM, [Online], https://www-01.ibm.com/software/info/bpm/what-is-bpm/ (Accessed : March 2019).

[5] P. Harmon, “Service oriented architectures and BPM,” (BPTrends), [Online] February 2005, http://www.bptrends.com/ (Accessed : June 2018).

[6] [11] C. Møller, R. Tan and C. Maack, “What is business process management? A two stage literature review of an emerging field,” Management Stream, 1(10), 9-13, 2009.

[7] VAN DEN BERGH, Joachim, THIJS, Sara, et VIAENE, Stijn. *Transforming Through Processes: Leading Voices on BPM, People and Technology*. Berlin : Springer, 2014.

[8] Harrington, H. James. *Business process improvement: The breakthrough strategy for total quality, productivity, and competitiveness*. McGraw Hill Professional, 1991.

[9] Harrington, H. James, et al. *Business process improvement workbook: documentation, analysis, design, and management of business process improvement*. McGraw Hill Professional, 1997.

[10] Hammer, Michael. "Reengineering work: don't automate, obliterate." *Harvard business review* 68.4 (1990): 104-112.

[11] S. Page, “The power of business process improvement: 10 simple steps to increase effectiveness, efficiency, and adaptability,” AMACOM Div American Mgmt Assn, 2010.

[12] A. Shtub and R. Karni. “ERP: the dynamics of supply chain and process management.” Springer Science & Business Media, 2010.

[13] Limam Mansar, S., and Hajo A. Reijers. "Best practices in business process redesign: use and impact." *Business Process Management Journal* 13.2 (2007): 193-213.

[14] Jansen-Vullers, M. H., et al. "Performance measures to evaluate the impact of best practices." *Proceedings of Workshops and Doctoral Consortium of the 19th International Conference on Advanced Information Systems Engineering (BPMDS workshop)*. Vol. 1. Trondheim: Tapir Academic Press, 2007.

[15] N. Brand, H. Van der Kolk, “Workflow Analysis and Design, Kluwer Bedrijfswetenschappen,” Deventer, 1995 (in Dutch).

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr

[16] S.W. van Cranenbroek, “CHANGING BUSINESS PROCESSES,” Eindhoven, July 2009.

[17] Alaaeddine YOUSFI, et al. “Smart Business Process Improvement for Process-Aware Information Systems,” 2015.

[18] Gordon Masiuk, “Developing and using process performance metrics”, Calagary Community of Business Process Management Professionals, June 2012.

[19] Kayode Ashogbon, “ Business Process Improvement through Knowledge management,” 2012.

[20] Ned Kock, “Business Process Improvement through E-Collaboration: Knowledge Sharing Through the Use of Virtual Groups,” 2005.

[21] Ismaili-Alaoui, A., Benali, K., Baïna, K., & Baïna, J. (2018, April). Business Process Instances Scheduling with Human Resources Based on Event Priority Determination. In *International Conference on Big Data, Cloud and Applications* (pp. 118-130). Springer, Cham.

[22] Ismaili-Alaoui, A., Baïna, K., Benali, K., & Baïna, J. (2018, June). Towards Smart Incident Management Under Human Resource Constraints for an IoT-BPM Hybrid Architecture. In *International Conference on Web Services* (pp. 457-471). Springer, Cham.

[23] Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... & Burattin, A. (2011, August). Process mining manifesto. In *International Conference on Business Process Management* (pp. 169-194). Springer, Berlin, Heidelberg.

[24] P. Russom, “Big Data Analytics,” [Online] 2011, ftp://ftp.software.ibm.com/software/tw/Defining_Big_Data_through_3V_v.pdf (Accessed : March 2019).

[25] D. Opolon, T. Saleh , E. Lenhard and N. Gordon, “How to Avoid the Big Bad Data Trap,” [Online] 2015, https://www.bcgperspectives.com/content/articles/big-data-digital-economy-how-to-avoid-big-data-trap/ (Accessed: March 2019).

[26] Wickham, H., & Grolemund, G. (2016). *R for data science: import, tidy, transform, visualize, and model data*. " O'Reilly Media, Inc.".

[27] Steven Engelhardt, [Online], 2017, https://www.stevenengelhardt.com/2017/12/01/visualizing-latency-part-3-rendering-event-data/, (Accessed: March 2019).

[28] Janssenswillen, G., Swennen, M., Depaire, B., Jans, M., & Vanhoof, K. Enabling Event-data Analysis in R.

[29] Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... & Burattin, A. (2011, August). Process mining manifesto. In *International Conference on Business Process Management* (pp. 169-194). Springer, Berlin, Heidelberg.

[30] Business Process Analysis in R, [Online], https://www.bupar.net/index.html, (Accessed: March 2019).

[31] Silvereco. Silver économie. http://www.silvereco.fr/

[32] rviews, “what is the tidyverse”, https://rviews.rstudio.com/2017/06/08/what-is-the-tidyverse/ (Accessed: March 2019).

[33] Janssenswillen, Gert, and Benoît Depaire. "BupaR: business process analysis in R." (2017).

*Corresponding author
Email address: abir.ismaili-alaoui@loria.fr