Data Analysis to Better Understand Business Process Models Discovered with Process Mining

Alessandro Bettacchi, Alberto Polzonetti, Barbara Re
Computer Science Division
University of Camerino, Camerino, Italy
e-mail: {*name.surname*}@unicam.it

Knut Hinkelmann
School of Business
FHNW, Olten, Switzerland
e-mail: knut.hinkelmann@fhnw.ch

***Abstract*—Due to continuous changes in the business context, enterprises have to rapidly react to novel market scenarios. To this end, a better understanding of the actual business processes is needed. This was the real need of a manufacturing company producing coffee machines. As-is processes have been investigated to understand in detail how the production chain works. First, we applied process mining techniques which produced models fitting the expectations, but also presenting some deviations from the designed flow of production activity. In order to understand the reason behind such deviations, an in-depth data analysis using On-Line Analytical Processing has been performed. Such awareness allows the management board to re-organize the production process. We also generalized the approach by proposing a methodology that allows to define, and potentially improve, the production, by giving recommendations.**

Keywords–Smart Manufacturing; Process Mining; OLAP.

I. INTRODUCTION

In the globalized market, the continuous changes in the business context, the increasing customer demands and shorter product life-cycle determine a highly competitive environment that forces manufacturing companies to a continuous alignment of the production and the internal organization. Research in the area of smart manufacturing tries to give an answer to such emerging needs. Indeed, smart manufacturing can be defined as "the dramatically intensified and pervasive application of networked information-based technologies throughout the manufacturing and supply chain enterprise" [1]. In particular, due to the complexity of the manufacturing production processes, a deep understanding of the as-is processes is essential to be able to quickly adapt such processes to new scenarios. This also enables a continuous improvement of the production, preventing bottlenecks, avoiding unexpected behaviors, and minimizing workarounds enforced by the workers. A deep understanding of the as-is scenario was the real need of the manufacturing company, producing professional coffee machines, that motivated our study. After several meetings, we agreed with the management board that an in-depth investigation of the production process is mandatory to continuously improve the way to work. This means learning from the past to better perform in the future.

To this aim, we analyzed the as-is production process using process mining techniques [2]: we applied five algorithms that we evaluated according to quality criteria [3] and complexity metrics [4]. The Inductive Miner (IM) alghoritm [5] proved to be the most suitable for the case under study [6]. Together with the production manager, we assessed the discovered process models detecting several unexpected behaviors. This finding prompted the need to understand the issues of such behaviors.

In this work, we use data analysis to investigate the factors influencing the process context [7][8] of the discovered process models. To be more precise, we focus on the correlations between the unexpected process behaviors and the context information. We selected meaningful data from the information system and generated a Data Mart (DM). Then, by the use of On-Line Analytical Processing (OLAP) tools, we outlined several correlations that we reported to the production manager for a better evaluation. As a result, we detected additional issues to consider for enhancing the production process and we provided some recommendations to prevent exceptional behaviors.

Based on the results achieved in the case study, and inspired by the process of Knowledge Discovery in Databases (KDD) [9], we also propose a novel methodology, named Process Deviations Causes Discovery (PDCD). PDCD relies on two main pillars: (i) *Process mining* for discovering as-is process models; and (ii) *Data Warehousing* and *OLAP* for analyzing correlations between the behaviors observed in the mined process and external events. The main aim of PDCD methodology is to achieve a greater awareness of unexpected behaviors detected in discovered process models. Particularly, after mining the as-is processes, the methodology allows to investigate the external factors, namely the context, affecting unexpected behaviors and to provide recommendations for improvements.

The remainder of this paper is organized as follows. Section 2 shows a motivating case study, Section 3 reports the data analysis activity, while Section 4 details the PDCD methodology. Section 5 presents some results coming from the implementation of the methodology. Section 6 provides related works and, finally, Section 7 concludes the paper.

II. UNDERSTANDING MANUFACTURING PROCESS

Here, we use a case study on a manufacturing company. The company produces professional coffee machines, which are exported all over the world. The manufacturing consists of assembling components provided in most of the cases by external suppliers. The production process is spread over six production lines numbered from 1 to 6. Each production line is organized into stations, each with a specific task and identified by letters from A to F. According to the different types of coffee machines, the organization of the stations in the production lines may vary: lines 5 and 6 have only five stations because B and C are merged and their activities are performed together.

- Station A assembles the frame of a new coffee machine and activates the Radio-Frequency IDentification (RFID) tag associated to it.
- Station B handles the hydraulic system. In the production lines 1 to 4, this station only assembles a portion of the hydraulic system, while in lines 5 and 6 the entire hydraulic system is assembled.
- Station C finalizes the assembly of the hydraulic system (this station is not relevant for lines 5 and 6).
- Station D deals with the electrical circuit.
- Station E performs the testing on several coffee machines simultaneously.
- Station F completes the coffee machine production including the packaging.

The company is assisted by a customized Information Technology (IT) system for managing the production process and all the related activities, such as production planning, reorder point, warehouse management, workers' support in all production phases, etc. The IT system, named ASCCO, is implemented as Process-Aware Information System (PAIS) [10]. The system also deals with tracking all the information related to the production line (assembly steps and times, faults, fixes, etc.) that are recorded in event logs. We extracted more than 450,000 events related to six years of production of 32 different coffee machines models. We then executed process mining on such event logs using five different algorithms, and we evaluated the results according to specific metrics concluding that the IM algorithm is especially suited for the case under study. These activities have been extensively discussed in our previous work [6].

The process models discovered with process mining showed some behaviors that deviate from the standard production process: the production manager and his staff were partly able to interpret such models and to react accordingly by reorganizing some phases of the production process. Despite that, the production manager required further investigation to explore special "events" that could affect the non-standard behaviors detected. We focused on production plans, workers, fixes and faults detected, customizations and the execution times of stations activities.

III. DATA MART IMPLEMENTATION

In order to make an efficient and comprehensive analysis, we rely on a Data Warehouse (DW) [11]. We started defining the conceptual model according to the Dimensional Fact Model (DFM) notation [12], as shown in Fig. 1, then we proceeded modeling the corresponding star schema.

The investigated fact refers to any single activity (summarized with the letter of the corresponding station) for assembling a coffee machine. The only measure is the execution time, that denotes the time required to perform each single activity, because most of the analysis relies on the *COUNT* operator for counting the number of items, as performed activities or produced coffee machines, according to the considered dimensions and their combination or aggregation. The dimensions we adopted are the date of assembly, the coffee machine model, the production plan, the engaged operator, faults and fixes accomplished, and the sequence of events (trace) generated by the comprehensive assembly. The dimensions were organized into appropriate hierarchies for enabling different levels of data granularity.

We implemented the Data Mart as a single cube, then we started the analysis through OLAP tools. We executed the interactive analysis involving stakeholders to take advantage of their domain knowledge and insights. The analytical tools allowed us to infer interesting considerations.

A. *Findings from OLAP Analysis*

We performed OLAP analysis with SpagoBI [13], a Web-based open source suite for business intelligence. The user interface allows a lot of processing, but in some cases we needed to modify the MultiDimensional eXpressions (MDX) query, auto-generated by the tool, in order to insert commands not available in the interface. An outcome of OLAP analysis concerned non-standard production sequences: their trend over time for any coffee machine model is proportional to the number of coffee machines of the same model produced in the same period. This means that non-standard behaviors are not related to special periods, influenced by some specific event, but they only depend on the production progress.

1) Low Execution Time and Unusual Production Line. Analysis results highlighted two unconventional situations at a glance: a large number of stations with low execution times (even less than a minute), as shown in Table 1, and a considerable number of specific coffee machines models produced on lines 5 or 6, rather than on the lines 1 to 4 as specified in the PAIS. An example is presented in Fig. 2.

We evaluated such findings assisted by the production manager and we realized that specific models, during periods of overproduction, are also assembled in lines 5 and 6, by changing the assembly process, rather than in the lines 1 to 4 for which they are designed. Moreover, the low execution times depended on an uncommon use of the traceability system. This issue was not known so far, because the production manager prepares a report on production progress on a daily basis taking into account only the number and type of produced coffee machines regardless of assembly time.

2) Weaknesses of the Information System. The above findings revealed some weaknesses of the information system: the RFID manual reading often leads it to record non effective execution times, while the rigidity of the

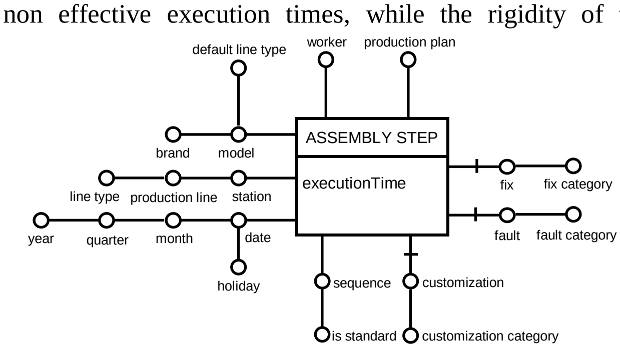


Figure 1. The ASSEMBLY STEP fact scheme.

production tracking system logs the C station activities, using fictitious times, also in the line 5 and 6, for all models designed to be assembled in lines 1 to 4. The management board decided to immediately implement improvements: (1) automating the RFID reading, in order to have absolute start and end time period of activities, (2) introducing more flexibility in the production tracking system and in particular for recording only activities really performed.

These improvements contribute to have a better event log files to be used for future process mining. In addition, tests on the new system for RFID automatic reading proved the effectiveness of such upgrade by logging inconsistent execution times for less than 0.1% of cases.

3) Customizations, Production Plans and Failures Effects. The data analysis on non-standard traces disclosed interesting connections. Many models showed an increase of customizations between 15% and 25% if compared to the global production of the same model. Similarly, the number of performed fixes and reported faults was well above the average values calculated for all the produced machines, in some cases even twice as much for the most common models as shown in Table 2. Furthermore, more than half of the traces were included in a few production plans. The above values do not seem a mere coincidence for coffee machines showing non-standard behaviors in the production process.

4) Knowledge Workers Activity. The assesment of non-standard traces revealed that a few workers seem to perform most of the activities, while in standard traces the workers are homogeneously distributed in stations. Such unexpected behavior, suggested the management board, is due to a few employees who are knowledge workers with a lot of experience, but who do not follow properly the procedures.

5) Information System Exceptions. The investigation also shows that a small part of non-standard traces comes from exceptions generated by the information system. Such traces should be marked to avoid noise in future analysis.

More generally, the observed results were thoroughly assessed by the production manager, who concluded that a good portion of non-standard behaviors was caused by operating procedures not compliant with company guidelines. These attitudes negatively affected the PAIS in recording sequences and timing of activities. The observed facts led the management board of the company to revise several aspects of the production process, and to request an upgrade of ASCCO to reflect such changes in addition to the two updates mentioned above. The planned improvements are presented next.

- New and enhanced operating procedures for the production process that will be properly fulfilled by the workers since exceptional behaviors will no longer be admitted.
- An enhanced alert system, integrated in ASCCO, for (i) reporting in real time exceptional behaviors in order to quickly react, and (ii) warning workers about previous faults encountered in the coffee machine model that they are assembling, in order to prevent the same issues.
- A new approach for performing critical customizations, consisting in specific procedures for assisting workers and different timing than the regular assembling.

IV. PDCD METHODOLOGY

The case study we run confirms how the use of OLAP analysis contributed to a better understanding of the discovered process models. This also contributed to achieve a better awareness and understanding that may be used to reorganize and improve the processes under study. It also helped to generalize the procedure we follow in a wider applicable methodology [14].

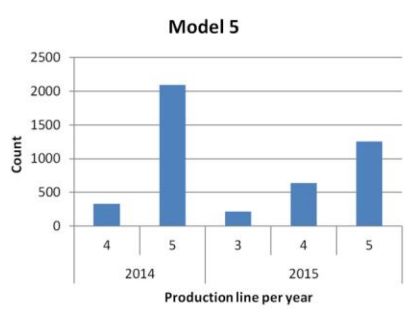
Here, inspired by the KDD process [9], we outline a methodology, named PDCD, which, starting from the selection of event logs, leads to improved process models in two steps. (i) *Process mining* for discovering as-is process models. (ii) *DW* and *OLAP* to analyze the correlations between the observed behaviors and external events.

Fig. 3 shows the basic flow of the proposed methodology. It is characterized by a high degree of interaction with the user, and it may require multiple iterations and present loops between some successive steps.

The availability of data is the beginning and the pillar of the approach. Such data may come from one or more information systems or may be gathered from many different sources such as spreadsheets, flat files, emails, etc. and then organized in a uniform and consistent manner.

TABLE I. SHORT EXECUTION TIMES

Station	Model					
	7	*8*	*12*	*15*	*18*	*19*
A	3.22%	9.08%	9.09%	2.79%	15.07%	16.98%
B	16.99%	12.63%	15.78%	8.96%	89.20%	86.67%
C	19.45%	24.80%	33.27%	38.06%	-	-
D	7.19%	13.95%	26.39%	4.47%	18.92%	21.80%
E	90.42%	86.70%	76.15%	57.83%	81.22%	85.77%



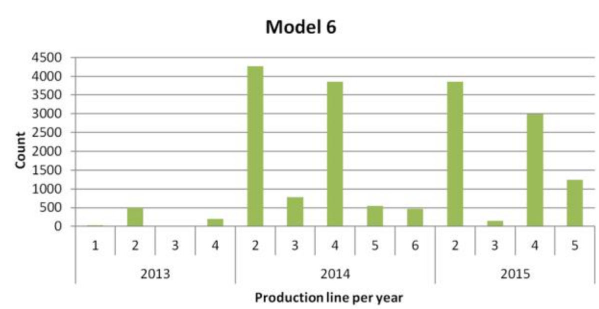


Figure 2. Two samples of coffee machine types assembled also in lines 5 and 6 even though they should be assembled on the lines from 1 to 4.

TABLE II. FIXES AND FAULTS DETECTED IN COFFEE MACHINES

Model	fixes (non-std.)	faults (non-std.)	overall fixes	overall faults
7	12.50%	10.79%	10.63%	8.93%
8	23.08%	19.23%	18.41%	8.94%
12	21.87%	21.88%	14.76%	15.24%
15	53.85%	38.46%	18.56%	16.40%
18	45.46%	50.00%	15.47%	13.87%
19	11.27%	11.26%	10.01%	8.96%

The first step, **Extraction**, consists of extracting suitable events, from the available data, as input for process mining. Events extraction means, firstly, to determine the appropriate information for the process, in order to produce an event log choosing only those data closely related to the scope of the analysis. This is necessary because, according to the adopted standpoint, it is possible to extract different event logs from the same data set. Event logs are usually stored in one of the typical formats: eXtensible Event Stream (XES) or Mining eXtensible Modeling Language (MXML).

The second step, **Process mining**, is applied to discover process models. It includes the choice of the process mining algorithm(s), the initial settings, such as parameter values, conditions or termination criteria, and the option to convert the resulting model into a different notation. The algorithms are generally chosen based on their characteristics and experiences performed in the same or similar domain. Sometimes, it is required to proceed in an empirical manner by applying several algorithms, and then determining which algorithm is the most suitable to the case under study using quality measures [3] and complexity metrics [4]. The most feasible process models will be used in the next steps.

The third step, **Evaluation**, is about evaluating the discovered process models: they are assessed and analyzed involving interested parties to understand the actual behavior of the system under study, and eventually comparing it with the desired behavior to focus on exceptions, and then making assumptions on the overall observed behavior.

The fourth step, **Data selection**, is about creating a target data set: the understanding of the actual process behavior and the assumptions made in the previous step, suggest the data subset to be selected for further investigation, among all the data initially available, in order to find probable connections to external causes related to the observed process.

The fifth step, **Preprocessing**, is data cleansing and transforming. It includes all the operations required to improve the quality of data selected in the fourth step, such as converting types and formats, removing duplicates, managing conflicts and inconsistencies, concatenating or separating relevant information, and defining methods for handling missing and unknown values. The outcome is a consistent, homogeneous and correct data set.

The sixth step, **Data modeling**, is building a Data Mart. It includes the conceptual design, for determining facts, measures and dimensions with related hierarchies, the logical modeling, for expressing the multidimensional model, e.g. the star or snowflake schema, the physical implementation, namely creating data structure according to the multidimensional model, and, at the end, the data feeding.

The seventh step, **OLAP analysis**, is finding correlations by using OLAP tools, with which to explore and analyze multidimensional data for outlining relationships between discovered process behavior and external factors impacting on the process, e.g. people involved, seasonal patterns, workload and resource availability, process misapplication. For this purpose, a domain expert is asked to actively interacts with such tools drilling-down, rolling-up, slicing and dicing, and pivoting, for generating several meaningful outcomes in form of (hierarchical) tabular data or charts for more friendly investigation and comparison.

The eighth step, **Interpretation**, is inferring the analysis results: specialists try to give a basis to the assumptions made, discarding those improbable, confirming the most probable, or requiring further investigation. This could lead to additional iterations returning to any of the previous steps.

The ninth and final step, **Discovered knowledge enactment**, is managing the new awareness on processes: implementing discovered process models in information system, if process-aware, or using such models to replace, or partially modify, the current ones, or simply using them as new reference models, for reorganizing the real processes to reflect new models, preventing the recurrence of specific deviations and exceptions, as well as providing guidelines and recommendations to process improvement.

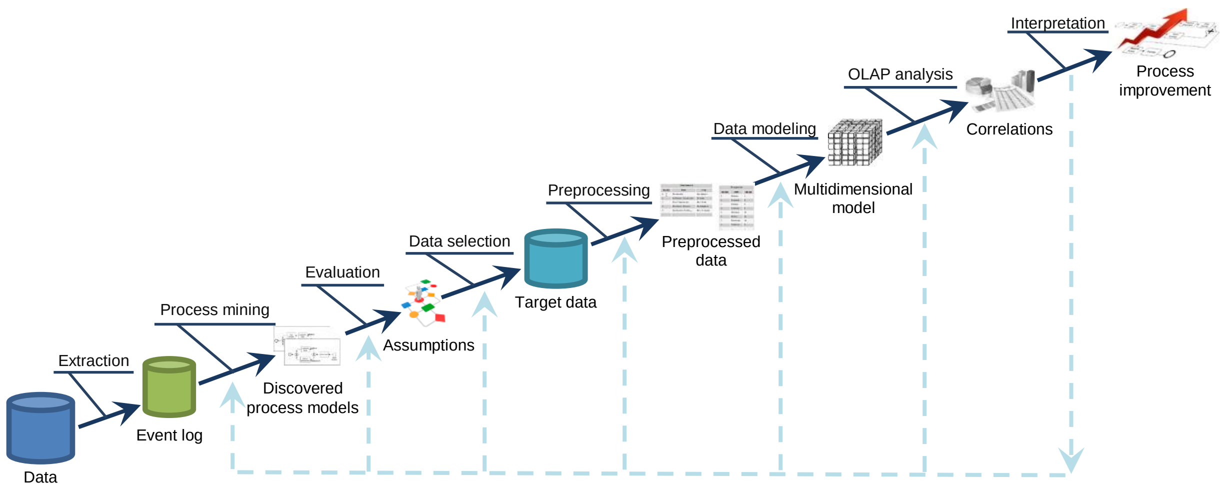


Figure 3. PDCD Methodology: Steps Overview.

V. Results and Discussion

The proposed methodology proves to be effective in the considered case study, leading to good results in getting a full knowledge of actual production processes and related context. The process mining activity alone allows to discover as-is processes, by providing models that ensure a close correspondence to the actual behavior of the processes because they are generated based on real event data. Therefore, the discovered models represent processes as they are actually performed during the examined period, but they did not provide any details about specific observed patterns.

One issue that often arises from stakeholders during the evaluation of a process model is “why this sequence of activities?”. The answer can rarely be inferred from the model itself. In our case study, it was not possible to answer such question even if the model presents a small number of activities. This became more and more complex in case involving a higher number of activities. In practice, it is only possible to make assumptions, which, however, must be validated in order to be “converted” into an answer.

The main goal of our methodology is precisely to try to give such answers. To this aim, the assumptions provided by the process experts are relevant for choosing the information to be investigated. This avoided to persist on irrelevant data or data not related with process under examination. The decision to address the DW world and to use OLAP tools revealed all the benefits in performing data analysis in a flexible and structured manner, observing information from different viewpoints and at different levels of detail. Furthermore, the data analysis phase cannot ignore the involvement of domain experts for attaining substantial results that will be further assessed by the same experts. The resulting suggestions could be used for:

- Accomplishing a new full cycle after generating more appropriate event logs;
- Repeating the analysis integrating the already used information, or using a different set of information;
- Establishing criteria to simulate changed processes, for checking the runs of processes before upgrading;
- Defining guidelines or take measures to improve processes and limit the exceptions.

The use of a systematic approach to provide criteria by which to argue the observed behavior in process models discovered by process mining, represents an added value to acquire a deeper understanding of the entire process. In addition, if the discovered model may be compared to a standard designed model, such criteria should support further assessment of detected deviations. Further assessment could determine which deviations to keep on the new process model and which ones to consider just as exceptions or, even, which ones to avoid because counterproductive.

VI. Related Work

To the best of our knowledge, there are no previous works that merge process mining and data analysis for discovering process models and then investigating external factors affecting such processes. The external factors are usually identified as the context of the process. In Business Process Management (BPM) the concept of context has several facets: in [8][15], it is described as the environment in which a business process may be used, while in [16], it is “The minimum set of variables containing all relevant information that impact the design and execution of a BP”, and in [17], the context is “any information reflecting changing circumstances during the modeling and the execution of a BP”. The work in [7] outlines the importance of considering the process context for improving BPs, learning from past experiences. In the above works, the concept of context is introduced to explain the benefits of the context-awareness in the BPM scope, and, in particular, in BPM design. However, also process mining techniques, as highlighted for the first time in [18], may greatly benefit in considering the process context, that is categorized into four classes: instance context, process context, social context and external context. In our work, we mainly considered the instance context, namely the factors that influence the singular process instances such as product customizations, assembling times and fixes.

The proposed methodology, inspired by KDD process [9], merges the BPM and DW. In particular, OLAP techniques are used to better understand the process models discovered by process mining, and the external causes, i.e., the context. A similar idea is in [19] where an approach for analyzing and preventing exceptions in BP is described. However, such approach is based only on generic Data Warehousing and data mining techniques. No process mining is applied and exceptional behaviors are defined by conditions over process execution data, i.e., subjectively, instead of comparing discovered models to standard ones.

In literature, there are further works that combine BPM and DW, with an extensive discussion being presented in [20]. The work in [21] describes a multidimensional approach for process modeling which enables the mapping of different aspects of a BP into a data cube and the support for a wide range of analysis. Such approach deals with only numerical data, a drawback overcome in [22] where the concept of OLAP data cube is merged with BP formalizing the notion of Process Cube. In Process Cube, events and process models are organized using different dimensions, and slice, dice, roll-up and drill-down operations are reformulated to be consistent with the new data structure. Improved versions of the Process Cube framework are the Process Mining Cube tool [23] and the PMCube explorer [24], that allow a more detailed multi-dimensional representation of a business process for a proper analysis.

Other works that combine BPM and DW, without merging them, show data mining as the main pillar. In [25] data mining is integrated with BPs for enabling an agile and exhaustive analysis of processes. Moreover, a methodology for successfully reusing data mining solutions during integration is introduced. Inspired by such approach, a formal framework for BP redesign is proposed in [26]: operational data collected during process runs are mined to explicitly represent the dynamics of the BP. This allows to re-design the process more efficiently. In [27] data mining is used to support decisions on resource allocation. Process context data, extracted from past process executions, are mined to acquire new knowledge for guiding optimal resource allocations in new process instances.

In summary, no work has so far combined process mining and DW as we did in this paper. They are strictly related for achieving a thorough knowledge of discovered processes by exploiting the context information, but they are not merged in order to quickly implement improvements in each single stage, such as a more efficient process mining algorithm or a better design of data cube.

VII. CONCLUSION AND FUTURE WORK

In a competitive and globalized business context, manufacturing companies need to adapt rapidly to new conditions in order to advance. Furthermore, production processes in the manufacturing field are quite complex, so it is needed to have a comprehensive understanding of such processes in order to adapt them to the new settings. In a case study, we investigated, in collaboration with domain experts, the process models discovered with process mining algorithms. We selected a large set of data from the company information system, we run process mining and we built a Data Mart. Using OLAP we then performed a thorough analysis and submitted the results to the production manager and his staff. Their interpretation contributed to a deeper understanding of the observed behavior and led to feedback on how to improve the production process.

We also generalized the approach by proposing a methodology for achieving a better awareness of the process models discovered with process mining.

REFERENCES

[1] J. Davis, T. Edgar, J. Porter, J. Bernaden, and M. Sarli, “Smart manufacturing, manufacturing intelligence and demand-dynamic performance,” Comput. Chem. Eng., vol. 47, pp. 145–156, 2012.

[2] IEEE Task Force on Process Mining, “Process Mining manifesto,” in Business Process Management Workshops, vol. 99 LNBIP, no. PART 1, pp. 169–194, 2012.

[3] W. M. P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer Berlin Heidelberg, 2011.

[4] J. Mendling, Metrics for Process Models: Empirical Foundations of Verification, Error Prediction, and Guidelines for Correctness, vol. 6. Springer Publishing Company, 2008.

[5] S. J. J. Leemans, D. Fahland, and W. M. P. van der Aalst, “Discovering block-structured process models from event logs containing infrequent behaviour,” in LNBIP, vol. 171, Springer International Publishing, pp. 66–78, 2014.

[6] A. Bettacchi, A. Polzonetti, and B. Re, “Understanding Production Chain Business Process Using Process Mining: A Case Study in the Manufacturing Scenario,” in Advanced Information Systems Engineering Workshops: CAiSE 2016 International Workshops Proceedings, pp. 193–203, 2016.

[7] K. Ploesser, M. Peleg, P. Soffer, M. Rosemann, and J. Recker, “Learning from Context to Improve Business Processes,” BPTrends, vol. 6, no. 1, pp. 1–7, 2009.

[8] United Nations Centre for Trade Facilitation and Electronic Business, “Core components technical specification - part 8 of the ebXML framework.” p. 113, 2003.

[9] U. M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, “From Data Mining to Knowledge Discovery: An Overview,” in Advances in Knowledge Discovery and Data Mining, USA: American Assoc. for Artificial Intelligence, pp. 1–34, 1996.

[10] M. Reichert and B. Weber, Enabling Flexibility in Process-Aware Information Systems: Challenges, Methods, Technologies. Springer Berlin Heidelberg, 2012.

[11] R. Kimball, The data warehouse toolkit. Wiley, 1996.

[12] M. Golfarelli, D. Maio, and S. Rizzi, “The Dimensional Fact Model: A Conceptual Model for Data Warehouses,” Int. J. Coop. Inf. Syst., vol. 7, no. 2–3, pp. 215–247, 1998.

[13] SpagoBI Homepage, https://www.spagobi.org/ 2016.10.31.

[14] B. Flyvbjerg, “Five Misunderstandings About Case-Study Research,” Qual. Inq., vol. 12, no. 2, pp. 219–245, 2006.

[15] M. Born, J. Kirchner, and J. P. Müller, “Context-driven Business Process Modelling,” Jt. Proc. 4th Int. Work. Technol. Context. Bus. Process Manag. TCoB 2009. Conjunction with ICEIS 2009., pp. 17–26, 2009.

[16] M. Rosemann, J. Recker, C. Flender, and P. Ansell, “Understanding Context-Awareness in Business Process Design,” Proc. 17th Australas. Conf. Inf. Syst., 2006.

[17] O. Saidani and S. Nurcan, “Context-awareness for adequate business process modelling,” in Third International Conference on Research Challenges in Information Science, RCIS 2009, Morocco, pp. 177–186, 2009.

[18] W. M. P. van der Aalst and S. Dustdar, “Process Mining Put Into Context,” IEEE Internet Comp., vol. 16, no. 1, pp. 82–86, 2012.

[19] D. Grigori, F. Casati, U. Dayal, and M.-C. Shan, “Improving Business Process Quality through Exception Understanding, Prediction, and Prevention,” in Proceedings of the 27th Intl. Conf. on Very Large Data Bases, Italy, pp. 159–168, 2001.

[20] M. K. Shahzad, “Improving Business Processes using Process-oriented Data Warehouse,” PhD dissertation, DSV, KTH, Stockolm, Sweden, 2012.

[21] S. Mansmann, T. Neumuth, and M. H. Scholl, “Multidimensional Data Modeling for Business Process Analysis,” in Proceedings of the 26th international conference on Conceptual modeling (ER’07), pp. 23–38, 2007.

[22] W. M. P. van der Aalst, “Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining,” Asia Pacific Bus. Process Manag. First Asia Pacific Conf. AP-BPM 2013, Sel. Pap., vol. 159, pp. 1–22, 2013.

[23] A. Bolt and W. M. P. van der Aalst, “Multidimensional Process Mining Using Process Cubes,” in BPMDS 2015, EMMSAD 2015, Held at CAiSE 2016 International Workshops Proceedings, pp. 102–116, 2015.

[24] T. Vogelgesang and H. J. Appelrath, “Multidimensional process mining with PMCube explorer,” in Proceedings of the BPM Demo Session 2015, vol. 1418, pp. 90–94, 2015.

[25] D. Wegener and S. Rüping, “On reusing data mining in business processes - A pattern-based approach,” in Business Process Management Workshops: BPM 2010, Revised Selected Papers, vol. 66 LNBIP, pp. 264–276, 2011.

[26] T.-M. Truong and L.-S. Lê, “Towards a Formal Framework for Business Process Re-Design Based on Data Mining,” in BPMDS 2016, EMMSAD 2016, Held at CAiSE 2016 International Workshops Proceedings, pp. 250–265, 2016.

[27] R. Sindhgatta, A. Ghose, and H. K. Dam, “Context-Aware Analysis of Past Process Executions to Aid Resource Allocation Decisions,” in Advanced Information Systems Engineering: CAiSE, Ljubljana, Slovenia. Proceedings, pp. 575–589, 2016.

 ISBN: 978-1-61208-536-4