Event Log Extraction for the Purpose of Process Mining: A Systematic Literature Review

Conference Paper · October 2019

CITATIONS
0

READS
124

5 authors, including:

Dusanka Dakic
University of Novi Sad
8 PUBLICATIONS **11** CITATIONS
SEE PROFILE

Darko Stefanović
University of Novi Sad
46 PUBLICATIONS **561** CITATIONS
SEE PROFILE

Teodora Lolić
University of Novi Sad
21 PUBLICATIONS **17** CITATIONS
SEE PROFILE

Dajana Narandzic
University of Novi Sad
4 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Event Log Extraction for the Purpose of Process Mining: A Systematic Literature Review

Dakic Dusanka, Stefanovic Darko, Lolic Teodora, Narandzic Dajana, Simeunovic Nenad

Abstract Process mining bridges the gap between process model analysis and data-oriented analysis, such as data mining, by enabling automated discovery of process models, comparison of existing process models with an event log of the same process and improvement of existing process models. Process mining prerequisite is an information system that supports and controls real-life business processes and consequently stores event data, such as messages, transactions, logs etc., as event logs in some type of a database. Event data is then extracted, filtered and loaded into process mining software, where certain type of process mining can be conducted. Process-aware information systems (PAIS), which assume an explicit notion of a case to correlate events of a process, provide such logs directly. However, many information systems that support execution of business processes are not explicitly process-aware and due to the variability of the event data sources, this phase of process mining is challenging and the most time-consuming. Consequently, various event log extraction techniques, approaches and tools are being developed, both specific and generic. To make a contribution to the issue, this paper presents a systematic literature review conducted with the aim to answer the questions about genericity of the approaches, applicability by non-experts and developed feasible tools.

1 Introduction

Since the process-oriented view on organizations, some enterprise information systems have adopted explicit process concepts and offer "generic modeling and enactment capabilities for structured business processes" [1]. These information systems store data about process execution in some form. Process mining assumes that it is possible to generate an event log from this data, such that "each event in the event log refers to an activity (i.e., a well-defined step in a process) and is related to a particular case (i.e., a process instance)" [2]. However, not all PAIS record event data in such way. Beside this traditional Workflow Management Systems (WMS) and Business Process Management (BPM) systems, other information systems such as Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) systems are rather data-centric and object-centric, meaning that "their processes operate on multiple interrelated business objects, each having their own case identifier, their own behavior, and interaction with each other" [10].

Dakic Dusanka, Trg Dositeja Obradovica 6, dakic.dusanka@uns.ac.rs
Stefanovic Darko, Trg Dositeja Obradovica, darko.stefanovic@uns.ac.rs
Lolic Teodora, Trg Dositeja Obradovica 6, teodora.lolic@uns.ac.rs
Narandzic Dajana, Trg Dositeja Obradovica 6, narandzic@uns.ac.rs
Simeunovic Nenad, Trg Dositeja Obradovica 6, nsimeun@uns.ac.rs

Information about business objects (e.g. documents in ERP) are scattered through various data tables [14, 19]. Therefore, extracting event logs from various data sources is the most challenging phase of process mining and can present an obstacle, especially for non-experts, i.e., business analysts that do not have programming knowledge. In the last decade, approaches and corresponding tools were developed to support and automate event log extraction. However, there is yet no generic approach that researchers and practitioners agree on and each approach still has limitations. In current literature, there is no consolidation of these approaches, their characteristics and challenges, making it difficult for practitioners to conduct this initial phase of process mining.

This paper presents the results of a systematic literature review on the event log extraction approaches in the last ten years, conducted based on the guidelines for performing systematic literature reviews in software engineering by Barbara Kitchenham [2], with the focus on: Event log extraction approaches that non-experts can apply; a possibility of a generic event log extraction approach, applicable to extracting data from every information system and feasible tools that support event log extraction.

The remainder of the paper is organized as follows. Section 2 presents basic concepts of process mining and event logs. Section 3 describes the systematic literature review methodology. Section 4 presents conducted planning the review phase of the systematic literature review. Section 5 presents conducting the review phase of the systematic literature review, where primary studies were selected and the review results presented. Section 6 discusses the systematic literature review results in detail and Section 7 concludes the paper and suggests future research.

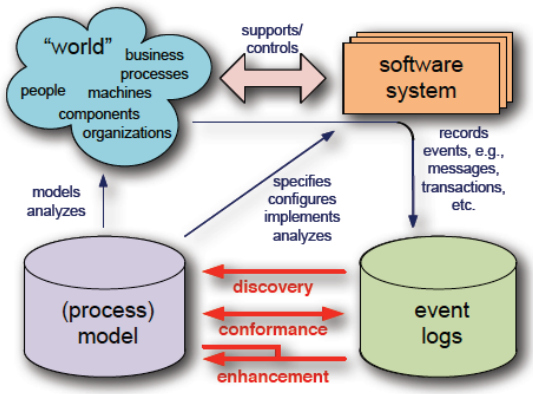
2 Process Mining and Event Log Concepts

Process mining is a new research field, introduced and defined by the IEEE Task Force on Process Mining in their Process Mining Manifesto [3]. They defined process mining as follows:

"The idea of process mining is to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs readily available in today's (information) systems. Process mining includes (automated) process discovery (i.e., extracting process models from an event log), conformance checking (i.e., monitoring deviations by comparing model and log), social network/organizational mining, automated construction of simulation models, model extension, model repair, case prediction, and history-based recommendations."

Fig. 1. presents prerequisites of process mining and a basic process mining procedure [3].

Fig. 1 Process mining procedure [3]



Software systems support and control real-life business processes, machines, components, organizations and people. Software systems are assumed to be configured and specified by process models. Consequently, these software systems

record events, e.g., messages, transactions, etc. that are later extracted from various data sources and constructed as event logs. Event logs are a starting point for performing three basic types of process mining: Discovery, conformance and enhancement.

According to the IEEE Task Force on Process Mining, in a constructed event log, it is assumed that [3]:

- "An event refers to a process activity or a task, which is a well-defined step in the process and is related to a particular case, i.e. process instance;
- Events are ordered;
- The case or process instance is a specific occurrence of a business process, while activity is an operation, part of a case, that is being executed;
- Each case has a unique identifier;
- An event log stores information about cases and activities, but also information about resources (person or device used for execution of the activity), event timestamps (moment when the event started and/or ended) or other data elements recorded with the event."

In 2010, the IEEE Task Force on Process Mining adopted XES (Extendable Event Stream) as a standard for process mining event logs [3]. XES is an XML-based format where every event log entry has an event type, a timestamp and other additional attributes.

The Process Mining Manifesto [3] outlined the guiding principles and main challenges of process mining. Guiding principles and challenges that relate to event logs and the scope of this literature review will be discussed. Guiding principles related to event logs are:

- GP1: Event Data Should Be Treated as First-Class Citizens;
- GP2: Log Extraction Should Be Driven by Questions;
- GP6: Process Mining Should Be a Continuous Process.

GP1 states that the quality of process mining is directly related to the quality of the event log, i.e. the input for process mining. Therefore, the IEEE Task Force on Process Mining suggests quality criteria that an event log should satisfy: "Events should be trustworthy, i.e., it should be safe to assume that the recorded events actually happened and that the attributes of events are correct. Event logs should be complete, i.e., given a particular scope, no events may be missing. Any recorded event should have well-defined semantics. Moreover, the event data should be safe in the sense that privacy and security concerns are addressed when recording the events."[1].

GP2 addresses the challenge of defining a scope of the process mining project, as data sources from which event data will be extracted contain various information, it is not trivial to decide what information to extract. Therefore, event log extraction should be driven by questions.

GP6 suggest that both historical event data and current data can be used for process mining in order to generate "living process models". Thus, an approach that enables continuous event log extraction should also be a prerequisite.

The following challenges are related to event logs:

- C1: Finding, Merging, and Cleaning Event Data;
- C2: Dealing with Complex Event Logs Having Diverse Characteristics;
- C10: Improving Usability for Non-experts;
- C11: Improving Understandability for Non-experts.

The main C1 challenges regarding finding, merging and cleaning event data are: "Data may be distributed over a variety of sources; Event data are often "object centric" rather than "process centric"; Event data may be incomplete; An event log

may contain outliers, i.e., exceptional behavior also referred to as noise; Logs may contain events at different levels of granularity; Events occur in a particular context (weather, workload, day of the week, etc.)" [3].

Challenges C10 and C11 both accentuate that current process mining techniques may be difficult to perform and understand for non-experts. Performance of process mining techniques includes event log extraction phase, hence this challenge applies on the event log extraction approaches as well.

3 Methodology

For the purpose of conducting this systematic literature review, a procedure for systematic reviews developed by Barbara Kitchenham [2] was followed. According to Kitchenham, a systematic literature review can be summarized into three main phases: Planning the Review, Conducting the Review and Reporting the Review.

3.1 Planning the Review

Planning the Review phase has a goal to determine and elaborate the need for a systematic literature review and to develop a review protocol. The need for a systematic literature review can be established by reviewing existing literature reviews in that particulair research area, following the guidelines proposed by Kitchenham [2]. Once the need for a review is established, a review protocol can be developed.

The review protocol shoud shortly present background and previously determined rationale for the survey. The second component of the review protocol is to formulate the research questions. A research question should, according to Kitchenham [2]: „be meaningful and important to practitioners and researchers, lead either to changes in current software engineering practice or increase confidence in the value of the current practice and identify discrepancies between commonly held beliefs and reality". Furthermore, question structure should take into consideration three different viewpoints: „population, i.e. group of people affected by the review, interventions, i.e. software technologies that adress specific issues and outcomes, i.e. factors of importance to practitioners".

After the research questions are formulated, a search strategy for primary studies based on initial scoping determines which search terms and databases will be used. Study selection criteria needs to be established, in order to be used as inclusion and exclusion criteria for initial analysis of primary studies and consequently, excluding studies from a systematic review.

The next step of the review protocol determines the quality assesment checklist that will be applied to the primary studies, included in the literature review. As Kitchenham states, there is no unique definition of"quality". However, other authors such as the CRD Guidelines [4] and the Cochrane Reviewers' Handbook [5] suggest that „quality relates to the extent to which the study minimises bias and maximises internal and external validity". By bias, a tendency to produce results that depart systematically from the 'true' results are considered. Internal validity is the extent to which the design and conduct of the study are likely to prevent systematic error. External validity, i.e. generalisability, applicability, is the extent to which the effects observed in the study are applicable outside of the study. The last steps are definition of strategy for the data extraction, synthesis of the extracted data and a project timetable.

3.2 Conducting and Reporting the Review phase

Once the protocol has been developed, the review can start. This involves the following activities [2]:
1. Identification of research;
2. Selection of studies;
3. Study quality assessment;
4. Data extraction and monitoring progress;
5. Data synthesis.

Reporting the Review is the final phase of a systematic literature review, where it is important to communicate the results of a systematic review effectively. Usually systematic reviews are reported in at least two formats [2]:

- In a technical report or in a section of a PhD thesis;
- In a journal or conference paper.

4 Planning the Review

First activity of Planning the Review phase should be elaboration of the need for a systematic literature review, by reviewing existing literature review of the subject [2]. However, there are no explicit systematic literature reviews on the subject of event log extraction, rather reviews of process mining field in general, its applicability, state of the art and challenges [6, 7, 8, 9].

Other sources of information about event log extraction approaches could be the related work chapters of studies that are pesenting a novelity on the subject. However, these stuides only mention previous event log extraction approaches relevant for their particulair topic of interest. This systematic literature review presents summarized information about current event log extraction approaches, their usability by non-experts, genericity and developed tools, by answering the research questions.

Based on the guidelines from Kitchenham [2], following research questions are formulated:

Q1: Are there any event log extraction approaches that non-experts can apply?

Q2: Is there a generic event log extraction approach, applicable for extracting event data from every type of information system?

Q3: Are there any feasible tools that support event log extraction?

For the purpose of this literature review, following databases where searched:

- SCOPUS,
- Web of Science and
- Google scholar.

Elsevier's Scopus is the largest abstract and citation database of peer-reviewed literature. Web of Science provides access to reliable, integrated and multidisciplinary research connected through linked content citation metrics from multiple sources within a single interface.

Search term defined for search in these databases is presented below:

“Process mining” AND “event log” AND extraction AND PUBYEAR > 2008

The inclusion criteria defined for this review are:

1. Paper has to present an approach for event log extraction for the purpose of process mining.

2. The approach has to include detailed information about developed procedure.
3. The study has to present a feasible approach, applicable in real-life scenarios.

Exclusion criteria defined for the review are:

4. Duplicate papers found in different databases should be removed.
5. If one author had more than one paper regarding the same approach, only one paper should be included in the review.

For the purpose of this literature review, the data extraction strategy was developed. For each primary study, following features will be extracted, in order to answer research questions:

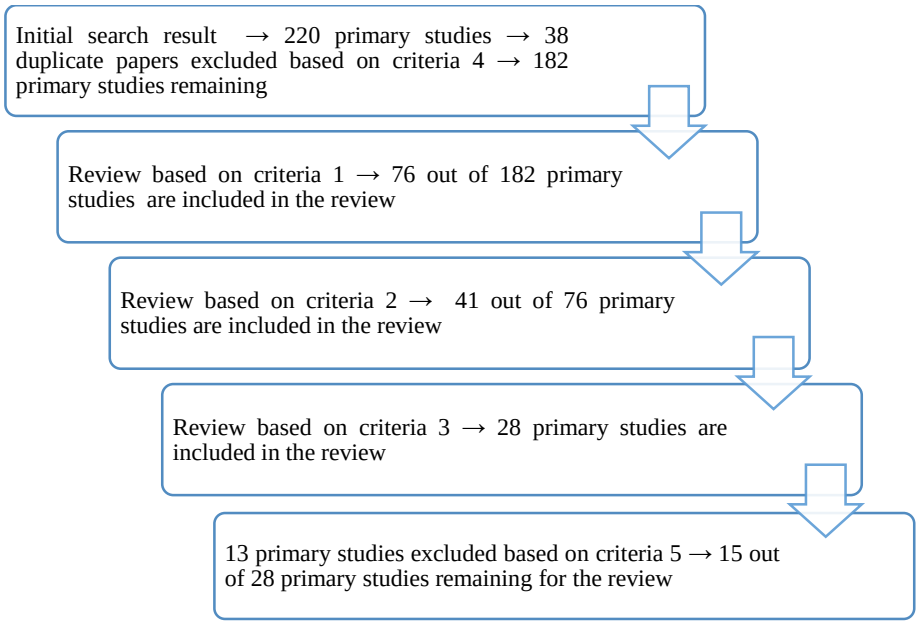
1. Publication year and source type,
2. Developed tool,
3. Feasibilty and applicabilty of the approach,
4. Genericity of the approach and information systems,
5. Applicability for non-experts.

5 Conducting the Review

5.1 Identification and selection of primary studies

First activity of conducting the review phase is identification of the research, i.e. primary studies that will be included in the systematic literature review. Primary studies are identified in accordance with inclusion and exclusion criteria, defined in the review protocol. Fig. 2 summarizes quantitative evidence of the inclusion/exclusion process in an appropriate flow diagram.

Fig. 2 Flow diagram of the exclusion and inclusion process



Search through data sources with previously defined search term resulted in identification of 220 primary studies. Based on exclusion criteria 4 (Duplicate papers found in different databases should be removed) 38 duplicate papers where found and excluded from further research. Review based on criteria 1 (Paper has to present an approach for event log extraction for the purpose of process mining) was conducted on the title and abstract of the papers and resulted in inclusion of 76 primary studies. Review based on criteria 2 (The paper has to include detailed information about the

developed approach) included 41 primary studies in the literature review. Out of 41 primary studies, 28 satisfied the criteria 3 (The study has to present a feasible approach, applicable in real-life scenarios). The last step of inclusion/exclusion process excluded 13 papers based on criteria 5 (If one author had more than one paper regarding the same approach, only one paper should be included in the review). Finally, 15 primary studies are included in the literature review.

5.2 Data extraction and summarization

In this chapter, main characteristic of primary studies are summarized and presented, following with the tables that present extracted data based on the data extraction strategy.

Fig. 3 presents event log extraction approaches by years in which primary studies that presented the approaches where published. The line diagram shows that two peaks in the number of published approaches occurred, one in 2012, two years after Process Mining Manifesto [3] was published and one in 2017.

Fig. 3 Summary of event log extraction approaches by years

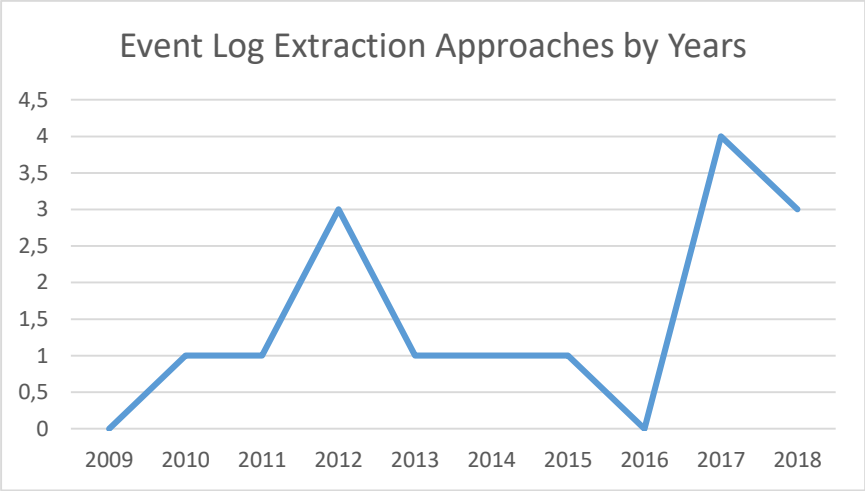


Table 1. presents source types of primary studies, showing that 40% of the approaches were published as journal articles, following with 27% of conference papers and master's thesis and one published procedure for event log extraction. Three out of four published master's thesis were conducted at the Eindhoven University of Technology.

Table 1 Source types of primary studies

Source type	Primary studies	%
Journal Article	10, 15, 16, 17, 19, 20	40%
Conference paper	12, 18, 22, 23	27%
Master's thesis	11, 13, 14, 24	27%
Procedure	21	7%

In the review protocol three research questions where formulated. Table 2. presents extracted data that answers the first question: Q1: Are there any event log extraction approaches that non-experts can apply?

Table 2 Primary studies and applicability of their approaches by non-experts

Applicable for non-experts?	Primary studies	%
yes	14, 16, 21	20%
no	10, 11, 12, 13, 15,17, 18, 19, 20, 22, 23, 24	80%

Table 2. shows that 20% of the reviewed approaches could be applied by a non-expert, while 80% o the approaches require extensive programming knowledge.

Table 3. presents extracted data that answers the second question:

Q2: Is there a generic event log extraction approach, applicable for extracting data from every type of information system?

Table 3 Primary studies and information system types that their approaches deal with

Information system types	Primary studies	%
ERP	10, 12, 13, 19, 21, 22, 23, 24	53%
Potentially generic	14, 16, 17	20%
Cloud systems	11, 20	13%
Non-process aware information systems	15	7%
Unstructured business processes	18	7%

During the data extraction phase, it was concluded that the reviewed event log extraction approaches can be grouped by different information system types. ERP systems, which are object and data centric systems, where used the most. Three approaches [14, 16, 17] showed potential to be generic, i.e. applicable for every data source and information system type. Furthermore, two approaches dealth with the extraction of event logs from cloud systems [11, 20], one approach presented a solution for event log extraction from completely non-process aware information systems [15] and lastly, [18] developed an approach for event log extraction from information systems that support unstructured business processes.

Table 4. presents information about tools and plug-ins that where developed in order to automate the approaches of event log extraction, sorted ascending by the year in which they were developed and with additional information about the event log format that they geneate as a result of event log extraction process. The third question (Q3: Are there any feasible tools that support event log extraction?) adresses the issue of the feasability of these tools. For each tool or plug-in developed, in corresponding primary study at least one case study was presented in order to demonstrate the applicability of the approach. Therefore, altough all tools and approaches have certain limitations, they are considered to be feasible.

Table 4 Developed tools and plug-ins with event log format

Tools/plug-ins	Primary studies	Year	Event log format
XES Mapper	14	2010	XES and MXML
Prototype for event log extraction from SAP ECC 6.0	13	2011	CSV
Xtract	12	2012	XES
Event-traces Injector (ETI)	15	2012	MXML
Eventifier	17	2012	XES
MANA	18	2013	XES
Xtract V2	10	2015	XES

Xtract V3	11	2017	XES
Ontology-driven extraction plug-in for ProM	23	2017	XES
Process mining tool and ERP connector	24	2017	XES
OpenSlex and PADAS	16	2018	new format, convertible to XES
XOC Log Generator Plugin in ProM 6	22	2018	eXtensible Object-Centric event logs (XOC)

6 Discussion

This section discusses the results of the conducted systematic literature review, presented in section 5.

The highest number of the reviewed approaches where published in the years 2012 and 2017. The previously mentioned Process Mining Manifesto was published in 2012. by the IEEE Task Force on Process Mining, hence the high number of approaches published that year. By the year 2017 these first event log extraction approaches where utilizied in the inudustry and showed limitations and challenges, as well as reasons for improvement. The approaches developed in 2017 and after, tend to solve this challenges. Most of the approaches were published in journal articles. However, a significant number of approaches were developed through master's thesis at the Eindhoven University of Technology, from where the idea of process mining originated.

For the purpose of answering the research question Q1, approaches that tend to improve usabilty of event log extraction by non-experts were identified. J.C.A.M. Buijs states in his master's thesis [14] that there is no tool that guides a business analyst towards converting data from a data source to an event log format suitable for process mining without the need to program and therefore creates a tool XES Mapper (later called XESame[1]). In [16] the authors recognized the need for a meta model that will integrate process and data perspectives and enable multi-perspective event-log building and analysis without the knowledge of SQL. In [21] the authors developed a procedure that leads a process analysts with limited knowledge of process mining though event log extraction process, increases the analyst's understanding of the decisions and their consequences, related to the choice of process instance, activities and attributes and provides the process analyst with a practical example or sufficient background information in order to conduct the event log extraction approach.

The question Q2 is answered positively, as there are three approaches [14, 16, 17] that can potentially be applied for every type of information system. The primary studies that reported the approaches presented case studies in which event log extraction is conducted on data sources of different information systems or at least provided a guideline for a generic application of their event log extraction approaches and tools. Two of these stuides [14, 16] where also found to be applicable for non-experts.

On the other hand, most approaches were developed for event log extraction from ERP systems [10, 12, 13, 19, 21, 22, 23, 24]. ERP system are object-oriented, i.e. along business documents, and not along business processes. To find the process data behind the static data, the business documents and related activities need to be matched into cases, i.e. the same execution of a process. Therefore, a suitable way of case identification has to be found which is also able to handle complex cases. Two main problems were identified while performing event log extraction from ERP systems: convergence and divergence [10, 11, 12, 13, 14, 22, 24]. In [14] the author defines them as follows: Convergence occurrs when the same activity is executed on

[1] http://www.processmining.org/xesame/start

multiple process instances at onc; Divergence occurrs when for one process instance the same activity is performed multiple times. Divergence and covergence remain unsolved problems, altough several primary studies that developed artifact-centric approaches suggested how to avoid these problems in particulair situations [10, 11, 14, 24].

As question Q3 is considered, various tools and plug-ins where developed to automate the process of event log extraction, as presented in table 4.

7 Conclusion

With todays highly automated business environment, amount of recorded data is rapidly growing. Correspondingly, new research disciplines that exploit this data, such as process mining, are emerging. Being a relatively new research discipline, process mining encounters challenges. This paper presented a systematic literature review of the most challenging phase of process mining – event log extraction. Existing approaches are identified and the following research questions are answered: Q1: Are there any event log extraction approaches that non-experts can apply? Q2: Is there a generic event log extraction approach, applicable for extracting event data from every type of information system? Q3: Are there any feasible tools that support event log extraction?

It is concluded that several approaches do provide higher understandability and usability by non-experts, however this is still a small number of approaches. Furthermore, some approaches do have a potential to be generic, i.e. applicable for every information system type. As feasible tools are considered, there are many developed protoypes and plug-ins that researchers can use to automate the event log extraction phase of process mining.

Future work should focus on detected challenges, such as convergence and divergence, continous event log extraction, improving usability for non-experts and present possible soutions.

ACKNOWLEDGMENT

This article has been produced as part of a research project: no. 47028 “Advancing Serbia's competitiveness in the EU accession process” supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia for the period 2011th-2019th year.

8 References

1. Van der Aalst, W. M. & Weijters, A. J. M. M. (2004) Process mining: a research agenda, Comput. Ind., doi: 10.1016/j.compind.2003.10.001
2. Kitchenham B. (2004), Procedures for Undertaking Systematic Reviews, In: Joint Technical Report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. (0400011T.1)
3. Van der Aalst, W. M. et al (2012), Process Mining Manifesto, In: BPM 2011 Business Process Management Workshops, Lecture Notes in Business Information Processing, Vol 99., Daniel F., Barkaoui K., Dustdar S. (eds),pp. 169-194, doi: 10.1007/978-3-642-28108-2_19
4. Khan, Khalid, S., ter Riet, Gerben., Glanville, Julia., Sowden, Amanda, J. and Kleijnen, Jo. (eds) Undertaking Systematic Review of Research on Effectiveness, CRD’s Guidance for those Carrying Out or Commissioning Reviews, CRD Report Number 4 (2nd Edition), NHS Centre for Reviews and Dissemination, University of York, IBSN 1 900640 20 1, March 2001
5. Alderson P., Green S, Higgins JPT. (eds) Cochrane Reviewers’ Handbook 4.2.2 [updated March 2004]. In: The Cochrane Library, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd

6. R'bigui, H. & Cho, C. (2017) The state-of-the-art of business process mining challenges, Int. J. of Bus. Process. Integr. and Man., doi: 10.1504/IJBPIM.2017.088819

7. Dakic, D., Stefanovic, D., Cosic, I., Lolic, T. & Medojevic, M. (2018) Business process mining application: a literature review, In: Proceedings of the 29th DAAAM International Symposium, pp.0866-0875, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-20-4, ISSN 1726-9679, Vienna, Austria doi: 10.2507/29th.daaam.proceedings.125

8. Tiwari, A., Turner, C. J. & Majeed, B. (2008) A review of business process mining: state-of-the-art and future trends, Bus. Process. Man. J., doi: 10.1108/14637150810849373

9. Van Der Aalst, W.M. (2012) Process mining: Overview and opportunities, ACM Trans. on Man. Inf. Syst., doi:10.1145/2229156.2229157

10. Lu, Xixi & Nagelkerke, Marijn & van de Wiel, Dennis & Fahland, Dirk. (2015) Discovering Interacting Artifacts from ERP Systems, IEEE Trans. on Serv. Comput., doi: 10.1109/TSC.2015.2474358

11. Santana Calvo, H.A. (2017) Artifact-Centric Log Extraction for Cloud Systems, Master's thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherland

12. H. J. Nooijen, Erik & F. van Dongen, Boudewijn & Fahland, Dirk (2012) Automatic Discovery of Data-Centric and Artifact-Centric Processes, In: Lecture Notes in Business Information Processing, doi: 10.1007/978-3-642-36285-9_36

13. Piessens, D.A.M. (2011) Event Log Extraction from SAP ECC 6.0, Master's thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherland

14. J.C.A.M. Buijs (2010) Mapping Data Sources to XES in a Generic Way, Master's thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherland

15. Ricardo Pérez-Castillo, Barbara Weber, Jakob Pinggera, Stefan Zugal, Ignacio García-Rodríguez de Guzmán & Mario Piattini (2011) Generating event logs from non-process-aware systems enabling business process mining, Enterpr. Inf. Syst., doi: 10.1080/17517575.2011.587545

16. González López de Murillas, E., Reijers, H.A. & van der Aalst, W.M.P. (2019) Connecting databases with process mining: a meta model and toolset, Soft. Syst. Model., doi: 10.1007/s10270-018-0664-7

17. Rodrı´guez C, Engel R, Kostoska G, Daniel F, Casati F, Aimar M (2012) Eventifier: extracting process execution logs from operational databases. In: Proceedings of the demonstration track of BPM 2012, vol. 940, pp. 17–22

18. Esposito, P. M., Vaz, M. A. A., Rodrigues, S. A., & De Souza, J. M. (2013). MANA: Identifying and mining unstructured business processes. In: Lecture Notes in Business Information Processing, doi: 10.1007/978-3-642-36285-9-20

19. Fliegner, W. (2014) Extracting process-related information from ERP systems for process discovery, Res. in Logist. & Prod., Vol. 4, No. 4, pp. 315-329

20. Bernardi, M. L., Cimitile, M., & Mercaldo, F. (2018) Cross-Organisational Process Mining in Cloud Environments, J. of Inf. & Knowl. Manag., doi: 10.1142/s0219649218500144

21. Jans, M. (2017) From relational database to valuable event logs for process mining purposes : a procedure, doi: 10.13140/RG.2.2.11343.69289

22. Li, Guangming & González López de Murillas, Eduardo & M. de Carvalho, Renata & Van der Aalst, Wil M. P. (2018) Extracting Object-Centric Event Logs to Support Process Mining on Databases, In: Information Systems in the Big Data Era, doi: 10.1007/978-3-319-92901-9_16

23. Calvanese D., Kalayci T.E., Montali M., Tinella S. (2017) Ontology-Based Data Access for Extracting Event Logs from Legacy Data: The onprom Tool and Methodology, In: Lecture Notes in Business Information Processing, doi: 10.1007/978-3-319-59336-4_16

24. Henny Selig (2017) Continuous Event Log Extraction for Process Mining, Degree project in information and communication technology, KTH Royal

Institute of Technology, School of Information and Communication Technology, Stockholm, Sweden