Expert Systems With Applications 133 (2019) 260–295

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review

# Process mining techniques and applications – A systematic mapping study

Cleiton dos Santos Garcia[a,*], Alex Meincheim[a], Elio Ribeiro Faria Junior[a,b], Marcelo Rosano Dallagassa[d], Denise Maria Vecino Sato[a,c,*], Deborah Ribeiro Carvalho[d], Eduardo Alves Portela Santos[e], Edson Emilio Scalabrin[a]

[a] Graduate Program in Computer Science (PPGIa), Pontifícia Universidade Católica do Paraná (PUCPR), Brazil
[b] Centro Universitário de União da Vitória (UNIUV), Brazil
[c] Instituto Federal do Paraná-Campus Curitiba (IFPR), Brazil
[d] Graduate Program on Health Technology (PPGTS), Pontifícia Universidade Católica do Paraná (PUCPR), Brazil
[e] Graduate Program in Production and Systems Engineering (PPGEPS), Pontifícia Universidade Católica do Paraná (PUCPR), Brazil

ARTICLE INFO

*Article history:*
Received 19 February 2019
Revised 6 May 2019
Accepted 7 May 2019
Available online 11 May 2019

*Keywords:*
Process mining
Workflow mining
Process mining applications
Process mining case studies

ABSTRACT

Process mining is a growing and promising study area focused on understanding processes and to help capture the more significant findings during real execution rather than, those methods that, only observed idealized process model. The objective of this article is to map the active research topics of process mining and their main publishers by country, periodicals, and conferences. We also extract the reported application studies and classify these by exploration domains or industry segments that are taking advantage of this technique. The applied research method was systematic mapping, which began with 3713 articles. After applying the exclusion criteria, 1278 articles were selected for review. In this article, an overview regarding process mining is presented, the main research topics are identified, followed by identification of the most applied process mining algorithms, and finally application domains among different business segments are reported on. It is possible to observe that the most active research topics are associated with the process discovery algorithms, followed by conformance checking, and architecture and tools improvements. In application domains, the segments with major case studies are healthcare followed by information and communication technology, manufacturing, education, finance, and logistics.



## 1. Introduction

Nowadays, there are registered events in many business processes, enterprise systems, automation and control systems, medical systems, daily activities, IoT devices, and social networks, among others. These events offer many possibilities to acquire knowledge to understand what is happening *de facto*, or which the most assiduous partners are. This initiative leads to the process mining area that is aimed to discover, check and improve real business processes from events available in many systems (van der Aalst, 2016). Traditional business process design starts from a detailed mapping approach involving multiple resources to establish a consensus model with the most recognized perspective of participants. As an alternative, process mining assumes that it is possible to obtain a meaningful process model extracting it from temporal documents or event logs readily available in system databases.

The process mining goal main challenges is to create a consistent and explicit process model given an event log and the use of tools to diagnose issues observing dynamic behavior (van der Aalst & Weijters, 2004). The identification of issues and diagnoses also needs explore the causal and casual (occasional) relations between activities, and this functionality is not present in a traditional Workflow Management System (WFMS) or Business Process Management System (BPMS). BPMS makes event log acquisition easier for process mining applications, but it is also viable to obtain event logs from different electronic transactions, registers, documents, or spreadsheets.

The Business Process Management (BPM) approach can be considered an evolution of Workflow Management. BPM is a structured and systematic approach to continuous process analysis.

* Corresponding authors.
*E-mail addresses:* cleitonsg@gmail.com (C.d.S. Garcia), alex.meincheim@hotmail.com (A. Meincheim), prof.elio@uniuv.edu.br (E.R. Faria Junior), mrdallagassa@gmail.com (M.R. Dallagassa), denise.sato@ifpr.edu.br (D.M.V. Sato), ribeiro.carvalho@pucpr.br (D.R. Carvalho), eduardo.portela@pucpr.br (E.A.P. Santos), scalabrin@ppgia.pucpr.br (E.E. Scalabrin).

https://doi.org/10.1016/j.eswa.2019.05.003
0957-4174/

According to van der Aalst and Damiani (2015), WFMS emerged in the mid-90s and focused on offering ways to automate some task integrated to a human task and to control the information flow. Some years later, BPMS became an extension of WFMS, which is more focused on operation analysis, management roles, and the work spread in the organization. However, the applications of WFMS or BPMS are very limited in many organizations owing to the difficulties in dealing with semi-structured or unstructured processes.

The concerns on how to identify process optimization and opportunities to achieve better results are continuously increasing; organizations are seeking to reduce time to achieve answers, reduce costs, maximize productivity, balance resource utilization, improve quality, minimize risk, and improve work well being. Business process re-engineering (BPR), Kaizen, Value Stream Mapping (VSM), Six Sigma, Lean Thinking, Value-Based Management (VBM), and Economic Value Added (EVA) have been adopted to improve efficiency and control business growth (Low, van der Aalst, ter Hofstede, Wynn, & Weerdt, 2017). Even though these approaches contribute to the performance improvement and reduction of costs, they have high failure rates reaching between 60% and 70% in BRP (Park & Kang, 2016). It is a motivation to seek more efficient approaches to process improvement and innovation.

In other side, for discovering patterns in large amounts of data, Data Mining (DM) has emerged. DM is based on techniques and methods focused on processing Big Data. The Big Data technologies, such as Hadoop, Hive, Impala, Spark and Storm, are mostly dedicated to processing large volumes of data for delivering traditional reports or dashboards, focusing specific activities (a slice of a business process) and rewriting the traditional BI (van der Aalst & Damiani, 2015). As Big Datas main research efforts involve storing and processing, other applications using analytical process identification and analysis of temporal event series patterns are not broadly studied in this area.

As an alternative for process improvement and aiming to fill the gap between DM and BPM, process mining has emerged (van der Aalst et al., 2012). The initial proposal of process mining or workflow mining was made in 1995, where the idea was to provide a set of techniques aimed to extract a process model based on recorded data in an information system. One of the first papers that opened the process mining discussion was by Cook and Wolf (1995, 1998), who proposed a process discovery approach using different methods from recurrent neural networks applied to probabilistic techniques and applications on software engineering data. In the same year, Agrawal, Gunopulos, and Leymann (1998) also analyzed previous approaches to sequential patterns mining and applied these in workflow scenarios. The intent is to provide support for automatically process models discovering based on data without investing large amounts of time on manual analysis. The initial process mining areas focus was to offer a new approach to business process design, which is complex and time-consuming, and the resulting process model usually is different from the process executed in reality (van der Aalst et al., 2003). According to Rozinat, de Jong, Gunther, and van der Aalst (2009), process mining basically observes the business activities using event logs and associated information to: 1) discover real process model; 2) perform a conformance check that compares the event logs to a defined process model; 3) enhance and expand the model combining performance, resource details, and bottleneck information, and so on (Rosa, Aalst, Dumas, & Milani, 2017).

Process mining does not replace the traditional process improvements methods such as Business Process Improvement (BPI), Continuous Process Improvement (CPI), Corporate Performance Management (CPM), Total Quality Management (TQM), Six Sigma, and others; however, process mining is able to enable, integrate, and accelerate process improvements. Several discovery process miners are able to accelerate process comprehensibility, modeling, and re-design (van der Aalst, Weijters, & Maruster, 2002; van der Aalst, de Medeiros, & Weijters, 2005; De Smedt, De Weerdt, & Vanthienen, 2014; Günther & van der Aalst, 2007; Jansen-Vullers, van der Aalst, & Rosemann, 2006; Lamma, Mello, Montali, Riguzzi, & Storari, 2007; Leemans, Fahland, & van der Aalst, 2013; Weijters & van der Aalst, 2003).

Process mining defines methods of distilling structured or understandable process based on data persisted during business activities (van der Aalst, Weijters, & Maruster, 2004). In this context, tools to improve, enhance, and audit the business processes were proposed. According to van der Aalst et al. (2012), process mining is capable to support compliance processes (e.g.: Sarbanes-Oxley). Process mining arose as a new link to Business Process Analysis (BPA) and Business Activity Monitoring (BAM) by applying DM to automatically detect real patterns and models (Reijers, Vanderfeesten, & van der Aalst, 2016).

The main goal of this systematic mapping is to provide an overview regarding process mining, following the identification of relevant areas of contribution in process mining, as well as the most discussed algorithms, and, finally, reporting application domains among different business segments. The systematic mapping study provides a comprehensive overview distilling the state-of-the-art from existing journals and conferences publications. This secondary study is helpful to researchers and practitioners looking for process mining information on relevant contributions organized according identified sub-areas, and it facilitates a broad view on hundreds of case studies and empirical applications in different business segments.

The article is organized as follows: In Section 2, the research method for systematic mapping is presented with the scope delimitation protocol. Section 3 outlines the main process mining research topics, establishing an overview of each one. The most relevant process discovery algorithms are presented in Section 4. Section 5 presents a diversity of process mining applications among a variety of business segments. Finally, a conclusion is given in Section 6, including a directions for future research.

## 2. Research method

This section covers the applied research method in this work conducted to search, select, extract, classify, and analyze previous work to understand the state-of-art direction of the investigation related to our exploratory research questions.

This study employed systematic mapping to identify, organize, and understand the main contributions of the state-of-art relating to process mining techniques and applications. Systematic mapping studies are designed to provide a wide overview of a research area, to seek and organize existing contributions, and provide quantity measures on the previous work (Kitchenham & Charters, 2007). In this context, this systematic mapping method was appropriate for seeking broadly studied areas. Fig. 1 illustrates the systematic mapping process with the essential process steps.

According to Petersen, Feldt, Mujtaba, and Mattsson (2008), each process step has an outcome until the final systematic map is achieved, as described below:

- Definition of research questions (research scope): are designed to provide a general scope for the study.
- Conducting the search: the primary studies are identified by using search strings on several selected scientific databases.
- Screening of papers: relevant papers are selected by applying the appropriate inclusion and exclusion criteria.
- Key wording of abstracts: Keywords are identified and combined to seek for high-level understanding of the research topic, thereby generating an organized classification or facet schema.

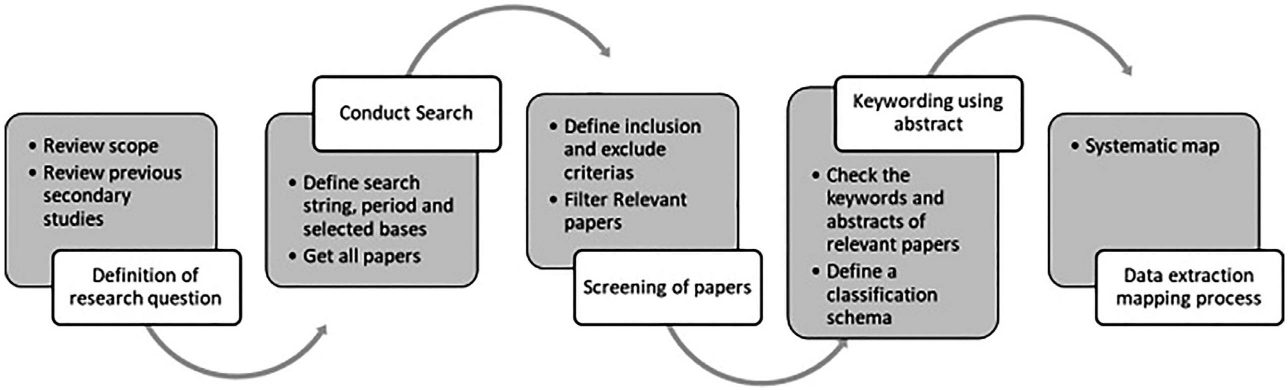


**Fig. 1.** Applied systematic mapping process, adapted from Petersen et al. (2008).

- Data extraction and mapping process: The classified papers in the previous step are extracted, aggregated, visualized, and mapped to answer our research questions.

### 2.1. Definition of research question

The goal of this paper is mapping and understanding the main contributions of the state-of-art toward process mining techniques and applications. Assuming this general objective, steps for conducting this research using questions were used to guide the Systematic Mapping Study.

- RQ1. Which research topics can be identified in the primary studies of process mining?
  This question aims to identify the process mining research topics and establish a categorized mapping from previous publications. It also maps the category of work experiencing growing attention and new emergent topics in the area.
- RQ2. What are the proposed process mining algorithms?
  The intent of this question is to highlight the different proposals for the specific area of process mining discovery. We also want to map the process discovery algorithms over time and identify the authors and the number of extended papers applying to each category of algorithms.
- RQ3. Where is process mining being applied?
  This question is aimed to understand the previous explored applications domain and/or industry segments. Initial searches regarding paper applications identified many publications in hospital or clinics pathways, perceived patterns steps of e-commerce customers, and manufacturing, among others.

As important as exploring the primary studies, this mapping first conducted an exploration in digital databases regarding secondary studies, seeking previous Systematic Literature Reviews (SLRs) and Systematic Mappings. The next paragraphs summarize these previous secondary studies between 2003 and 2018.

*2003 - Workflow mining: A survey of issues and approaches (van der Aalst et al., 2003).* Data & Knowledge Engineering. Five approaches to drive the process miner were explained to deal with different problems. The paper proposes a common XML format for event logs used to compare 4 existing process mining tools.

*2004 - Process mining: A research agenda (van der Aalst & Weijters, 2004).* Computers in Industry. Presents a general view, including 11 challenges to be addressed such as noise, incompleteness, loop, etc.

*2013 - Computer-interpretable clinical guidelines: A methodological review (Peleg, 2013).* Biomedical Informatics. Process mining is only one possible approach used to create computer-interpretable guidelines.

*2013 - The use of software product lines for business process management: A systematic literature review (dos Santos Rocha & Fantinato, 2013).* Information and Software Technology. Among 25 established attributes, only one focused on BPM Evaluation phases was aimed to extract primary works on process optimization using Business Activity Monitoring and Process Mining.

*2014 - Systematic review on security in Process-Aware Information Systems - Constitution, challenges, and future directions (Leitner & Rinderle-Ma, 2014).* Information and Software Technology. Discusses capacities to check compliance, identity inconsistencies, jump steps (fraud detection), root cause analysis in process mining to improve security.

*2015 - Compliance monitoring in business processes: Functionalities, application, and tool-support (Ly, Maggi, Montali, Rinderle-Ma, & van der Aalst, 2015).* Information Systems. Discusses compliance monitoring, defining main functionalities, and comparing approaches.

*2016 - The State of the Art of Business Process Management Research as Published in the BPM Conference (Recker & Mendling, 2015).* Business & Information Systems Engineering. The work is related to research methods, quality discussion, maturity, citations index, and progress in the business process management area.

*2016 - Process mining in healthcare: A literature review (Garcia, Ramirez, & Larrea, 2015).* Biomedical Informatics. The research covers 74 papers in healthcare focusing process mining usage.

*2016 - Process mining in oncology: A literature review (Kurniati, Johnson, Hogg, & Hall, 2016).* IC on Information Communication and Management, ICICM. This process mining study in oncology selected 37 papers from 758 articles.

*2018 - A systematic mapping study of process mining (Maita et al., 2018).* Enterprise Information Systems, EIS. This study analyzed 705 papers (2005–2014), mapping the relationship between data mining tasks in the process mining context.

Two surveys by van der Aalst et al. (2003) and van der Aalst and Weijters (2004) discussed the process mining area, sharing issues, techniques, and approaches. These studies helped to guide

**Table 1**
Digital libraries and search strings.

| Digital database | Search string |
|---|---|
| ACM Digital Library | "query": { content.ftsec:("process mining" "processes mining" "workflow mining" "mining workflow" "workflows mining" "mining workflows") }"filter": {"publicationYear":{ "gte":2002 }} |
| IEEE Xplore | ("process mining" OR "processes mining" OR "workflow mining" OR "mining workflow" OR "workflows mining" OR "mining workflows") &ranges = 2002_2018_Year |
| ScienceDirect | ("process mining" OR "processes mining" OR "workflow mining" OR "mining workflow" OR "workflows mining" OR "mining workflows") AND pub-date > 2001 |
| Springer Link | ("process mining" OR "processes mining" OR "workflow mining" OR "mining workflow" OR "workflows mining" OR "mining workflows") & facet-start-year = 2002 |

collaboration and support of new techniques and tools. These works were a starting step, showing many gaps for further contributions, algorithm improvements, and applications. As a secondary study, Peleg (2013) reviewed clinical practical guidelines, where process mining was identified as one solution for guidelines maintenance and compliance analysis. Another secondary study was focused on software product lines for business process management, and process mining was identified as an approach for process optimization (dos Santos Rocha & Fantinato, 2013). Process mining was identified as a prominent approach for security control in a systematic review on security (Leitner & Rinderle-Ma, 2014). A secondary study approached process mining focused on conformance monitoring, and established main functionalities and compared existing approaches and tools (Ly et al., 2015). A business process management study reviewed quality aspects of mapping evolution, maturity, impact, and indicated increasing community attention concerning process mining, discovery, or analysis (Recker & Mendling, 2015). Besides, two literature reviews focused on process mining in healthcare, first Rojas, Munoz-Gama, Sepúlveda, and Capurro (2016) mapped case studies according to 22 medical field areas, and Kurniati et al. (2016) reviewed papers focused on oncology applications. Finally, a systematic literature mapping Maita et al. (2018) identified primary studies on process mining from 2005 to 2014, focusing on how the data mining techniques and tasks have been applied to the process mining context.

Despite the existence of previous Systematic Literature Mapping or Reviews, few works were directly related to process mining techniques. Other open gaps on previous secondary work are related to establishing a broadly horizontal view of process mining sub-areas and on different business segments applications.

### 2.2. Conduct search

The general search string for this study was: ("process mining" OR "processes mining" OR "workflow mining" OR "mining workflow" OR "workflows mining" OR "mining workflows"). The chosen data sources were ACM Digital Library,[1] IEEE Xplore,[2] ScienceDirect[3] and Springer Link.[4] The filter period was defined from 2002 to cover the relevant discovery algorithm, called the Alpha algorithm or Alpha miner, until 2018. Table 1 presents the search string adapted to each digital library, and, it was used to consider all fields available in each database.

Table 2 lists the numbers of extracted papers from each database according our established search string. Unfortunately, at SpringerLink it was not possible extract all papers, as restrictions of our affiliated institution access.

**Table 2**
Numbers of papers by digital database in first search.

| Digital database | Number of papers |
|---|---|
| ACM Digital Library | 470 |
| IEEE Xplore | 1363 |
| ScienceDirect | 1172 |
| SpringerLink | 798 |
| **Total** | **3713** |

In the study selection process, we excluded articles based on abstract contents. In some cases, this was performed by reading the full text. Other studies were added as recommended using the snowballing. Each article was only reviewed by a single author. Inclusion and exclusion criteria was defined to filter papers for this study. The use inclusion (IC) criteria were:

- IC.1: The paper is electronically available and found by the search string in all fields and period;
- IC.2: At least two journal or conference reviewers;
- IC.3: Studies published online in the time frame 2002 to February 2018;
- IC.4: Snowball technique that includes some referenced papers not indexed in selected digital libraries, e.g.: Very Large Databases (VLDB) conferences, Eindhoven University workshops, and Society for Industrial and Applied Mathematics (SIM).

Subsequently the exclusion criteria (EC) were applied. These criteria were as follows:

- EC.1: Not written in English;
- EC.2: Duplicated work;
- EC.3: Not focused on process mining, e.g.: mining iron, metal mining, pollution, chemistry, or environment impact;
- EC.4: Referenced process mining only on introduction, or fundamentals, or part of the state-of-art. Not a process mining contribution or application;
- EC.5: Points to process mining only as further direction;
- EC.6: Only related to author biography;
- EC.7: Not a paper, such as only a contents guide, index, or marketing information;
- EC.8: Very short contribution, only one or two pages;
- EC.9: Doctoral Consortium or Standard proposals (ISO/IEC);
- EC.10: Unavailable papers or removed research retraction.

Fig. 2 shows the total of 3713 papers, 48 included papers by the snowballing, and the amount of excluded papers according each exclusion criteria. After applying all exclusion criteria, 1278 articles were selected for this systematic mapping.

### 2.3. Published papers and data extraction

The total of published papers by year was provided in Fig. 3. Process mining publications have increased year by year. In recent years, we can observe instability, but it is likely caused by the approval process for some papers. As this systematic mapping took two years to conclude, the first peak of publications was in 2014, and then the peak moved to 2015. This indicates that it is a growing knowledge area and the instability over the last two years is caused by the review and edition period, and by including only the early months in 2018.

Data extraction begins with generic information regarding the most active countries in Process Mining. These consist of The Netherlands, China, Germany, Italy, and the USA, as presented in Fig. 4. These countries contributions were identified based only

[1] http://dl.acm.org.
[2] http://ieeexplore.ieee.org.
[3] http://www.sciencedirect.com.
[4] http://link.springer.com/.

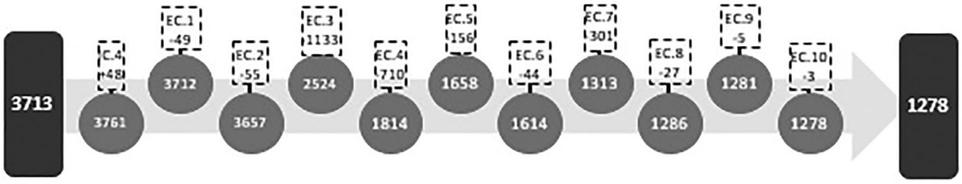


**Fig. 2.** Screening process by criteria.

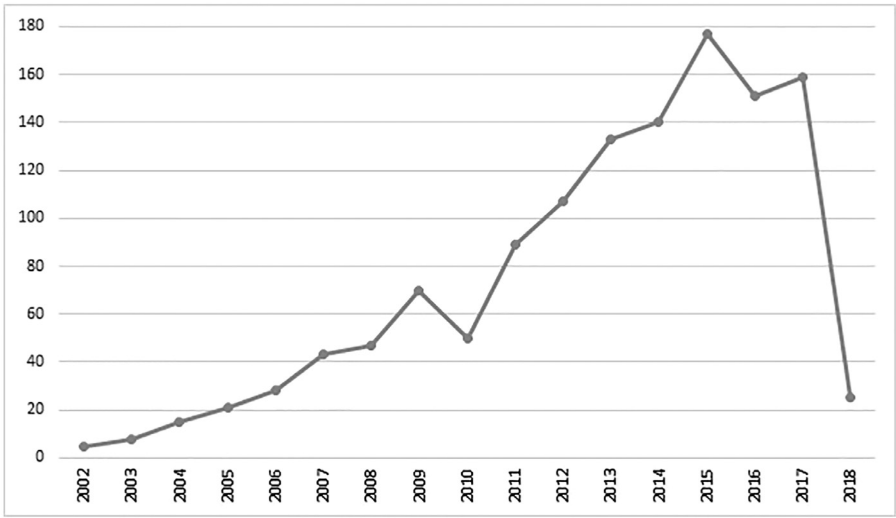


**Fig. 3.** Amount of publications on process mining by year.

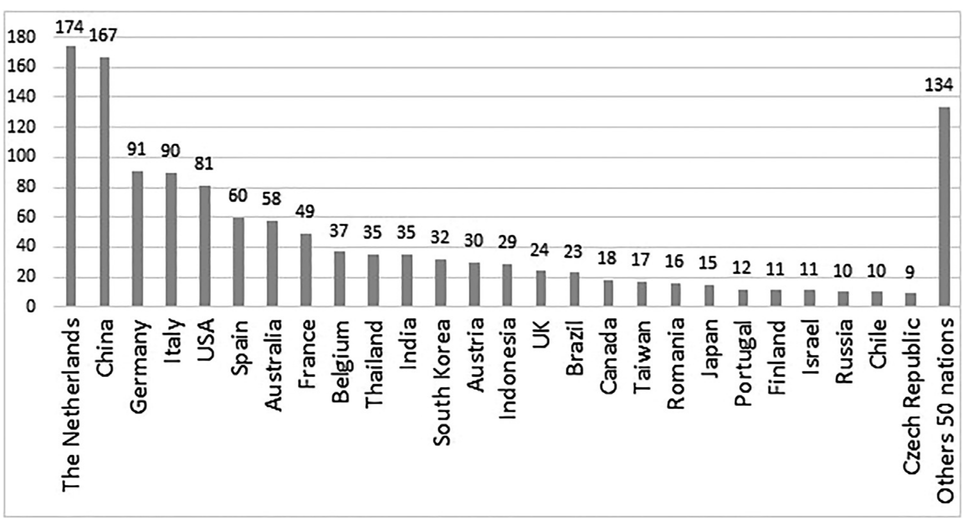


**Fig. 4.** Journals and conferences by country.

on first author affiliation. The Process and Data Science group from The Netherlands has produced significant contributions. For many years, Prof.dr.ir. van der Aalst led the group at Technische Universiteit Eindhoven (TU/e) and authored or co-authored more than nine hundred publications. Now he leads the Rheinisch-Westfälische Technische Hochschule (RWTH) in Aachen. Other relevant process mining research groups are present at: Katholieke Universiteit Leuven[5] from Belgium, Queensland University from Australia, Universität Ulm from Germany, UniversitÃ  della Calabria from Italy, Universität Innsbruck, and Vienna University of Technology from Austria. Significant research and contributions to process mining area have been achieved between partnerships involving more than one research group. As evidence, Prof.dr.ir. van der Aalst is co-author of approximately 29% of publications whose first author is affiliated with entities from other countries.

Other relevant information involves the journals and conferences that provided significant contributions. Journals contained 625 papers, and conference papers accounted for more than 603, followed by some book chapters. *Information System* is the journal that has more publications related to process mining, with a total of 47 papers, followed by *Expert Systems with Applications* with 31 papers, and *Decision Support Systems* with 27 papers. In the next section, we will outline the first research question.

[5] http://processmining.be/.

## 3. Process mining research topics

To answer our research question: "Which research topics can be identified in the primary studies of process mining?" all selected studies were analyzed and categorized according to highest contribution, as some papers included a secondary focus, we also identified a secondary categorization to the papers that have relevant contributions in two research topics, here also called categories. The systematic mapping explored the three main types (van der Aalst et al., 2012) of process mining: discovery, conformance and enhancement. Basically, the discovery type aims to mining process models by using process discovery algorithms (also called process miners); conformance focused in compliance checking and auditing; and enhancement aims to extend and improve processes, adding emphasis, information, and KPIs (key performance indicators). Besides this three main types, we defined more specific categories, based on the concepts described by van der Aalst (2016) and also by the first analysis of secondary papers. These categories were distributed between the three types of process mining according to the contribution of the papers. We also observed a relevant category related to the supporting areas for process mining projects. In this topic we included papers which are not directly describing one of the three main types of process mining, but explore other issues related to the implementation of process mining projects. We defined the following categories for this area which will be former described: Process Mining Applications, Methods for Process Mining Projects, Architecture and Tools, Background, Gathering and Cleaning and Ontology. Fig. 5 illustrates the identified categories including the three main types of process mining (splitted into categories) and the supporting areas topic.

Fig. 6 provides the distribution of papers by research topic.

The top ten research topics are responsible for 86.61% of contributions in process area. At 23.00% and 359 papers, the most active research topic is the process discovery algorithms, mainly aimed to understand activities control flow, since the first publications in 1998. The second active research topics focuses on conformance checking, reaching 12.75% and 199 papers. The third most relevant category is related to the process mining applications at 9.10%, followed by architecture improvements and tools at 8.36%, in fifth, papers focusing on methods for process mining projects, which defines practical and conceptual frameworks to conduct and apply process mining at 7.69%. Sixth is about clustering, abstraction, and process variant control proposals, representing 5.89%, followed by enhancement by extending process models at 5.45%. Finally, background knowledge papers with 5.25%; predictive techniques with 4.68% and organization or social miners, which include the organization perspective in the process models with 4.55% corresponding to the top ten research topics. The remaining seven categories represent only 13.39% of the total papers analyzed.

We could also observe on Fig. 7 the papers distribution between the three types of process mining and the supporting areas category. The category containing more papers is the supporting areas, this can be explained because in this topic we grouped papers related to projects, applications and tools. Between the three types of process mining, discovery represents the major number of papers at 480. Following discovery we can observe enhancement with 306 papers, which indicates that researchers aim to improve real business process.

In the next sections we will briefly describe the three main types of process mining and the identified research topics for each one. Besides, the supporting area will be presented with the identified categories.

### 3.1. Process discovery

This type of process mining is the precursor and the most explored type, where the analysis is mostly dedicated to a posteriori view of business process execution event logs (Weerdt, Backer, Vanthienen, & Baesens, 2012). The process miner is the usual term used for the algorithm capable of extracting the knowledge from event logs and creating a process model. This discovery process should deliver a model with quality, where it is expected to be comprehensible by avoiding unnecessary complexity, and deliver acceptable accuracy, balancing recall, precision, and generalization (Weerdt et al., 2012).

According to de Medeiros and van der Aalst (2008), process discovery can support managers to easily answer questions, e.g.: how are the process instances (i.e. cases) being executed, what are the highest frequent paths, how is the frequency distribution according each process paths?

Previous studies report challenges related to low efficiency of process discovery applied in scenarios with high noise processes, repetitive activities, different types, and variants of processes combined in the same logs (Yang & Su, 2014). This study was applied in the healthcare area, but the contribution can be generalized to other areas, the work by Yang and Su (2014) points to future studies divided into four problems: analysis and identification of process variabilities; flow customization with adjustments according to the contextual situation, resource availability, and probable results; integration with process management; and automated self-learning improvements or recommendations.

Despite process discovery being very effective in capturing process control behavior, there are limitations in the process models supplementary discovery related to identifying responsible roles, input or output documents, and identifying and generalizing decision points (Akman & Demirörs, 2009). In Section 4, the most common process discovery algorithms will be presented.

The next topics will describe each research topic identified in the mapping study directly related to the process discovery task.

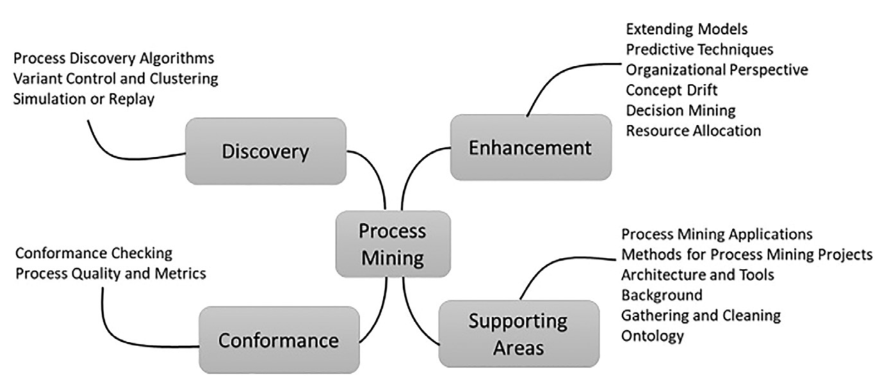


**Fig. 5.** Process mining identified categories.

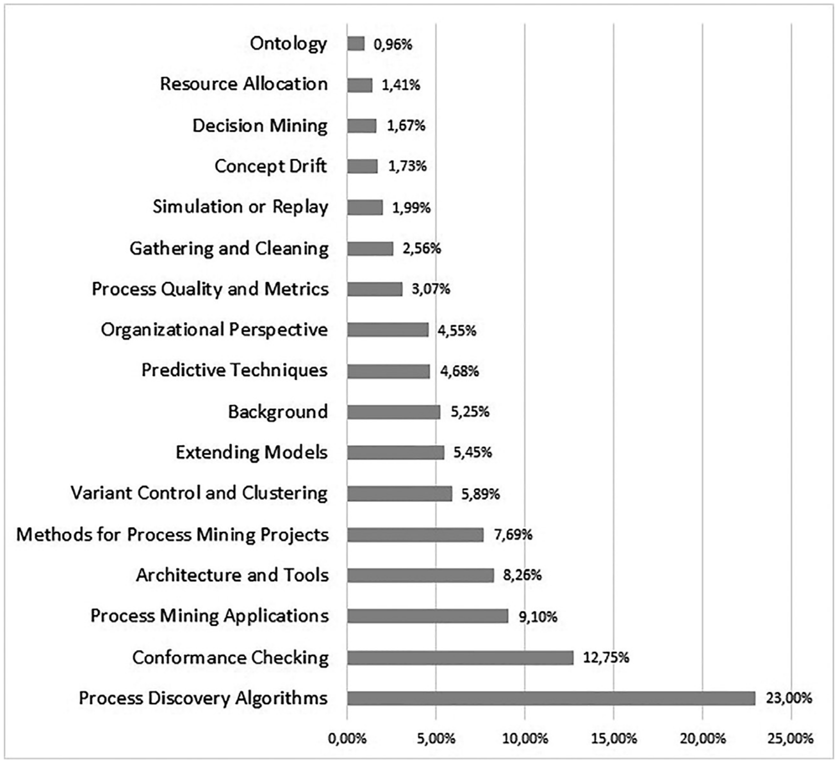


**Fig. 6.** Detailed distribution of papers by process mining research topics.

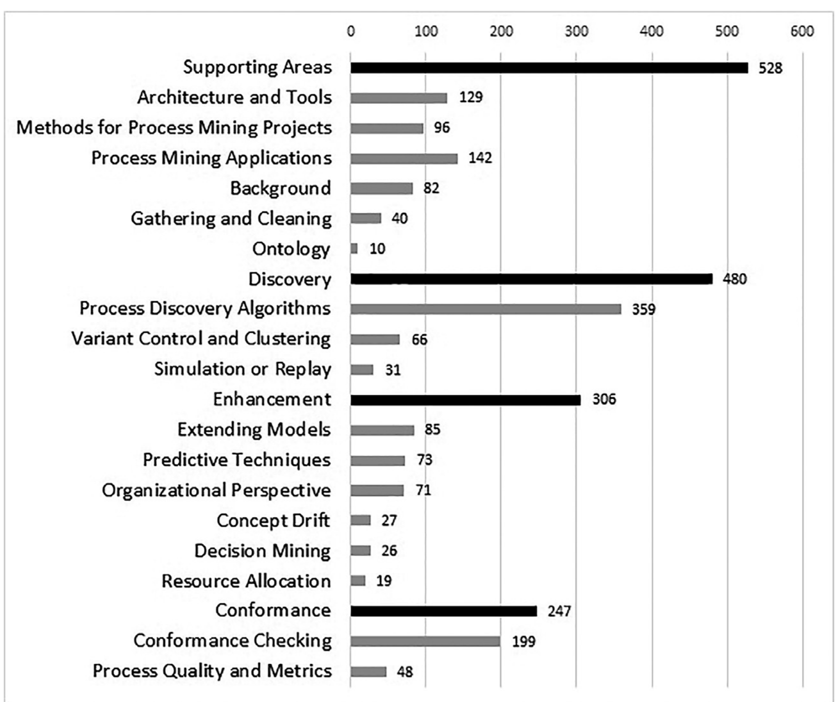


**Fig. 7.** Distribution of papers by main categories.

Even though conformance checking and enhancement make use of the process discovery techniques to obtain the process model to check or enhance, the following topics are concerned only on the discovery process.

*3.1.1. Process discovery algorithms*

Many process discovery algorithms have been proposed, such as the alpha algorithm, heuristic mining, multiphase mining, fuzzy mining, genetic mining, region miner, integer linear programming (ILP) miner, declarative miner, etc. Previous comparisons reported that some algorithms involve high computational cost and several approaches are very difficult or unfeasible to apply in industry (Wang, Wong, Ding, Guo, & Wen, 2013).

The process discovery algorithm is responsible to generate a process model based on any a-priori information, such as an event log (van der Aalst, 2016). The process miners need an event log containing ordered events indicating that some activity happened. Also, the events must be related to a case identifier, which indicates a single instance of the process, e.g. a person paying for a product bought online. The ordering of the events is essential to identify the causal dependencies in process models. The process discovery algorithms will be explored in more detail on Section 4, which aim to detail the algorithm developed over time.

#### *3.1.2. Variant control and clustering*

The process discovery algorithms aim to generate quality process model. The result of a process discovery algorithm can be evaluated considering four quality dimensions can be characterized: fitness, simplicity, precision, and generalization (van der Aalst, 2016); which will be further described. A process model with high precision, can be overfitting, thus resulting in a very complex and unstructured process, e.g., a spaghetti model. To avoid this, the process discovery algorithms can use clustering and variant control techniques.

As in conformance analysis, clustering requires a matching condition, such as: lexical applying syntactic or/and semantic similarity measures on events or task labels; structural matching using process topology; and, behavioral matching considering process semantics.

Aiming to deal with this complexity, one of the first approaches proposing the idea of the clustering process was Günther and van der Aalst (2007). This work aimed to balance the abstraction and aggregation, applying three steps: preserve highly significant behaviors, aggregate the less significant behaviors with highly correlated behaviors, and, remove the less significant and lowly correlated behaviors. However, this work aims to extract one overall view of an unstructured process, without considering the enhancement visualization of each relevant process variant. This first combination of clustering and process miners is available in the fuzzy miner implementation of the ProM Framework. Process mining can easily acquire insights related to a group of process variants, where we can check specific process variant according to frequency, execution time, or other relevant process KPI guiding the process improvement.

Focus on extracting clustering according to similar process variants, many of the existing clustering techniques do not ponder the process models quality during the clustering procedure, which significantly impacts the clustering bias (Weerdt, vanden Broucke, Vanthienen, & Baesens, 2013). The majority of process discovery algorithms are focused on generating fitness or precision models with simplicity; however, distinct behaviors that generate separate models is more relevant for easier analysis, diagnostics, and improvements (de San Pedro & Cortadella, 2016). The ActiTraC algorithm was proposed by Weerdt, vanden Broucke et al. (2013) considering the measurements in the sampling. It is classified as unsupervised, and is divided in phases: selection, look ahead analysis, and residual paths evaluation. First, a selective sampling strategy was employed based mainly on frequency measurements.

A recent paper by Prodel, Augusto, Jouaneton, Lamarsalle, and Xie (2018) proposed method capable to deal with complex traces. The most significant contributions were to start with a compact hierarchical and flexible representation of event classes delivering meaningful aggregations; subsequently, the replayability was improved using Monte Carlo sampling method. This approach allows decreasing the computational complexity (event-log-size), and, finally for process model optimization the *tabu* search algorithm seeks promising neighborhoods until the stop condition is met. The method was tested on a wide range of both generated and real event logs. This approach outperforms the state-of-the-art process mining algorithm, including the commercial tool DISCO™. The Prodel et al. (2018) proposal achieved relevant results combining hierarchical aggregation and an optimization algorithm to balance the model fidelity and complexity.

*Process abstraction, aggregation, and variant clustering.* Process abstraction is an overview of a process, for example, focusing on the most frequent or relevant effort, time consumption, or requirements and their execution costs. Usually, this technique is implemented by hiding the activities, but it is also possible to use aggregation. Aggregation aims to group related activities to make the process more compact and easier to understand. This feature is very well applied in maps and GPS applications. Other very relevant capabilities are the clustering of process variants, where such variants are divided by affinity, which improves the diagnostics according to the characteristics common to the different clusters (Li, Reichert, & Wombacher, 2010).

*Clusters based on vector space.* Clustering log traces allows identifying the different process variants, as Greco, Guzzo, Pontieri, and Sacca (2006) proposed, it is possible to group similar execution flows, thus minimizing the number of process models and turning each model into a more concise version. This approach allows defining hierarchical patterns for refinement of each model, and it is possible to combine this pre-analysis step with any process miner algorithm, which is available in ProM.

Other clustering work was conducted by Song, Yang, Siadat, and Pechenizkiy (2013), who address the challenge of dealing with big event logs, especially unstructured and computational problems faced by the miners and the advantage of cluster techniques to reduce complexity. The study applies dimensionality reduction procedures, such as singular value decomposition (SVD), random projection (RP), and principal components analysis (PCA). The results accomplished are a good fitness with low computational cost. The authors compare three different clustering algorithms, the most commons in the DM field, such as K-means clustering, agglomerative hierarchical clustering and self-organizing maps.

Hompes, Buijs, van der Aalst, Dixit, and Buurman (2015) proposed an extended Markov cluster (MCL) algorithm able to discover different process groups with different sizes and densities. The proposal delivers a clustering algorithm with low computational costs and scalability for graphs. This technique is capable of handling different control-flows in the variants as well as in the outliers, making it possible to explain both common and exceptional behaviors. A known issue for measures for vector similarity is the impact in order that is lost, but it was solved by using order in perspectives with occurrence of frequent patterns and also showing deviation from the mainstream flow. The clustering implementation was compared with two other implementations “Trace Alignment” and “ActiTraC”, and is available in ProM.

*Model-based sequence clustering.* Ferreira, Zacarias, Malheiros, and Ferreira (2007) and Rebuge and Ferreira (2012), inspired by the work of Cadez, Heckerman, Meek, Smyth, and White (2000), proposed a combination of first-order Markov models and expectation-maximization approach. The main contributions are related to a method executing a sequence clustering algorithm. It builds a diagram for cluster analysis, used to understand, first, regular behavior, and is then used for process variants and infrequent behaviors. Finally, the method includes an analysis of hierarchical sequence clustering, and selecting clusters for further study. This proposed method was applied in clinical event logs from a hospital in Portugal.

#### *3.1.3. Simulation or replay*

Replay aims to simulate all cases based on event log and given or discovered process model. It allows to observe each log trace and to show the logic between activities and transitions in the model, from the initial mark to the end or last step of each instance. It also allows to measure discrepancies between model and event logs being a useful instrument for behavior and conformance analysis (Rozinat, Mans, Song, & van der Aalst, 2009). Simulation can also be extended for automatically defining models to prediction of delivery capacity, adjusting parameters such as volume of process instances, and adding or removing resources.

### 3.2. Process conformance

Process Conformance is the type of process mining responsible to measure the quality of a process model. The quality of a process model is usually described considering the four quality dimensions (van der Aalst, 2016):

- Fitness: ability to observe the event log behavior in the discovered model;
- Precision: quality to avoid behavior unrelated to the used event log in the discovery process, like avoiding underfitting concept;
- Generalization: capacity to accept new similar events related to previous events used for discovering, like avoiding overfitting concept; and
- Simplicity: quality to be as simple as possible.

Process conformance is not only restricted to define the quality dimensions and how to measure them, but concerns also to the different applications of the conformance checking. Process monitoring based on conformance checks, for instance, is also a relevant and prominent topic. Some mistakes are usual when referring to conformance and compliance; these verification analyses are similar, but are not homogeneous concepts. The main difference involves when the process occurs: the conformance check is based on post mortem analysis with complete event logs (usually in an off-line mode), while compliance monitoring is based on partial and runtime event logs (Ly et al., 2015).

#### 3.2.1. Conformance checking

Conformance checking can be used to confront the model with reality acquired through the event logs on systems. According to van der Aalst (2012a), the verification of conformance can be used to verify the accuracy of the documented processes, point out different cases, and try to realize what they may have in common, and identify points of deviations in the process. In addition, for auditing purposes, verification of conformance can be used to calculate the efficiency of a discovered process model and improve a new or existing model. Conformance checking is used on several occasions, making it one of the pillars of process mining.

de Medeiros and van der Aalst (2008) also established the main "c-level" interest related to process conformance, e.g.: how compliant are the performed cases with the standardized and documented process models, what are the main issues, how frequent are the anomalies, and where are the business rules violated?

A previous systematic review performed by Ly, Maggi, Montali, Rinderle-Ma, and van der Aalst (2013) identified 10 relevant functionalities in the process conformance: 1) time metrics; 2) case data; 3) resource information; 4) explicit/implicit conditions associated with multiple events allowing different orders; 5) supporting activity lifecycle; 6) multiple-instance constraints e.g. time, resources, or other data; 7) reactively identifying and resolving compliance violations; 8) proactively detecting and managing violations; 9) explaining the root cause; and, 10) quantifying noncompliance using metrics. Additionally, Ly et al. (2013) conducted an evaluation involving five conformance proposals (ECE Rules, Mobucon EC, Mobucon LTL, SeaFlows, and Supervisory Control Theory) based on these 10 relevant functionalities.

Despite all of these proposals, according Association of Certified Fraud Examiners, the time for detection and for uncovering a disbursement scheme had a median of 13 months in 2016. However, the median duration of an undetected scheme is 18 months, and some frauds lasted at least two years before being detected (Association of Certified Fraud Examiners, 2016). This reality can change using online data solutions with process mining support.

#### 3.2.2. Process quality and similarity metrics

The discovered process and the applied filters should balance the quality perspectives: fitness, precision, generalization and simplicity. The most dominant observed metric is fitness that results in a value between 0 (absence of capacity for supporting all event traces) to 1 (perfect fitness). Three other quality dimensions were explored by Rozinat, de Jong et al. (2009), who proposed the appropriateness focusing on Occam's razor idea. This work also distinguishes the structural and behavior appropriateness, i.e., structural appropriateness by focusing on simple model that can support trace events (as a syntactic definition, avoiding duplicity), and, behavioral appropriateness, which tries to be too generic and allows for too many behaviors (how many process executions in the log are observed but are never used). Behavioral or structural similarities require a mapping approach to identify each activity in the models. To realize this, other techniques are usually combined to recognize a syntactic and linguistic similarity to identify each activity, for example, using string distance and synonyms (Dijkman, Dumas, van Dongen, Käärik, & Mendling, 2011).

In a recent systematic review, process similarities were organized into four dimensions (Schoknecht, Thaler, Fettke, Oberweis, & Laue, 2017): Natural language (for string edit distance, word edit distance, semantic similarity, and similarity of virtual bags optionally enhanced by natural language processing (NLP)-techniques); graph structure dimension (for similarity according edges, graph and tree edit distance, and feature-based similarity); behavior dimension (for longest usual sub-path, similarity of causal footprints, and similarity of behavioral profiles); and, human estimation dimension (for crowd-based similarity estimation, user feedback).

Some limitations are a consequence of the chosen process miner and the representation model, and among such different issues, van der Aalst (2016) defined: inability to represent concurrency (high complexity to realize); inability to represent silent actions or invisible transactions; inability to represent duplicate actions (simplicity and avoiding loops); inability to model OR-splits/joins (or complexity to combine AND and XOR structures to realize an OR operation using a Petri net); inability to represent non-free-choice behavior, that is a situation where the concurrency and choice meet (in other words, a mixture of choice and synchronization); and, inability to represent hierarchy (fuzzy miners define some clustering forms as activities or sublevel tasks).

The metrics are important to understand, measure, and balance the quality of a process model. As a similarity measurement for business process models, these metrics can support process mining techniques for: checking conformance; standardization or harmonization; searching for a model in repositories with a similar aspect; and, searching for adequate paths and reusing them automatically (Schoknecht et al., 2017).

### 3.3. Process enhancement

This type of process mining activity is focused on extending the process model with relevant information. For example, GPS mobile applications combine online information about traffic, giving emphasis to congested streets. Process mining can perform the same function by using timestamping in event logs, combining these with the process model for predictions using statistics or machine learning. This extended process models are very useful to provide operational support, which is the most ambitious form of process mining van der Aalst (2016).

The decision makers in organizations are very interested in the contributions of this perspective of process analysis because of the high aggregated business value and insight provided by process enhancements. According to de Medeiros and van der Aalst (2008), this type of process mining can support answers to questions such as: what is the average throughput time of cases, what are the transition and decision probabilities, which transitions are time-consuming, what are the critical activities or resources, what is activity duration (working time), and how much waiting time between activities?

A usual visual technique that allows observation of the process execution is called replay analysis. This technique allows to visualize all process instances within defined date time periods, performing each case representation over the process model and allowing time acceleration to obtain insights in a few minutes regarding what occurred over a period of weeks. According to van der Aalst (2012b), this establishes a well-defined meaning for all visual emphasis, such as:

- Activity size indicates the number of occurrence or metrics related to costs or resources usage.
- Activity color emphasizes the duration or working time.
- The width of a connection/arc can reflect the importance of each dependency.
- Connections color highlights waiting time for executing next activity, revealing bottlenecks.
- Activities position can show an implicit meaning, for example, identifying hub or support activities.

#### 3.3.1. Extending models

Even though many process discovery activities are concerned on the control-flow of a process model, it is possible to enrich the process model adding other perspectives by cross-correlating it with the log. This research topic include papers which extends the process models using additional information in the event logs. van der Aalst (2016) defines that we can obtain an integrated process model by adding different perspectives to it.

#### 3.3.2. Predictive techniques

Many predictive approaches have been combined with process mining, such as Bayesian inductors, decision trees, case based-reasoning, recommender systems, and neural networks, among other approaches. The goals of prediction are multiple, e.g., predicting instance duration, identifying delays, classification of root cause, predicting the next activity, and recommending a path of execution inside a process, among others.

A top-down Induction of Logical Decision Trees, called TILDE, was used by Vasilyev, Ferreira, and Iijima (2013) to predict the case duration. It is based on inductive logic programming (ILP), predicting the duration of a give case. This approach compares and separates cases into different groups according duration patterns to complete, it is an automatic grouping and it is not required any threshold definition. As result, a classification perspective based on ILP was capable of deducing the motive for each delay and the delay range. This knowledge was embedded in an event log, such as information regarding a user performing the task. From evaluation experiments, it was possible extract many rules, according users involved in the task and the estimated impact. All these reasons are inducted by the TILDE algorithm to identify delay impact.

Prediction of the completion instant of cases during execution was researched by Polato, Sperduti, Burattin, and de Leoni (2014). The prediction is based on two steps: predict the control-flow path using statistics (Naive Bayes), and, all relevant activity data are applied for forecast using a Support Vector Regressor to estimate the completion time. The technique was evaluated using a real-life case study achieving significant prediction results.

Prediction can also be used to improve process performance and mitigate risks in realtime, by taking proactive and corrective actions. Márquez-Chamorro, Resinas, Ruiz-Cortés, and Toro (2017) proposed an approach based on evolutionary algorithms for run-time prediction of business process. This approach was validated using two real-life datasets.

In general, predictive process monitoring techniques consider information obtained from the case (intra-case) whose measures of interest one wishes to predict. But in real life scenarios, the outcome of a running case may depend on all cases that are being executed concurrently. Senderovich, Francescomarino, and Maggi (2019) present a framework for feature encoding in predictive process monitoring considering two dimensions: intra-case dependencies and inter-case features to fill this gap.

A dynamic approach called PETRA exploited process events of a maintenance platform (Karray, Chebel-Morello, & Zerhouni, 2014). PETRA was divided in subsystems to track cases, induce learning, and obtain and reuse capitalized knowledge. Based on active learning techniques, the event log is processed for considering the dynamic situations using and accumulating context of process path, similar to an user experience. To organize the learning data, it uses maintenance domain ontology. The prediction approach is realized by a combined decision tree and rule mining to learn the maintenance factors. The simulation shows this approach is capable of reducing costs and can identify shortcuts according to current needs and with less effort.

Lakshmanan, Shamsi, Doganata, Unuvar, and Khalaf (2013) used decision tree learning methods to predict the next activity during the process execution and estimating the degree of certainty. The technique uses a discovered process model, calculate process probabilities, and extends the model using Markov chain to estimate the execution of all possible activities in the process, supporting parallel execution. The method was applied in automobile insurance validation.

Yang, Dong et al. (2017) proposed a recommendation system based on a regression model. This approach was divided into two stages. First, an analysis stage including: (1) clustering traces using a similarity measure for complex process traces based on timeline warping to seek the optimal pairwise alignment; (2) identifying a sample that represent each cluster; (3) regression model analysis to map correlations between cluster membership and any log extension data; and (4) visualizing models and statistical data. Then, in the second stage consist in: (1) predicting the cluster for a given process instance context, and (2) presenting recommendations based on clustering prototype.

Evermann, Rehse, and Fettke (2017) proposed a deep learning approach using recurrent neural networks to predict the next activity of a business process. This approach does not require the explicit process model, which turns the prediction independent of the quality trade-off of the models, and an alternative when the model is difficult to obtain. The event logs were translated into sequence of words to take benefits from the natural language processing. Two real datasets evaluated the approach showing interesting results.

#### 3.3.3. Organizational perspective

The research in this category aims to accomplish mining of organizational structures by combining social network analysis, mapping resource behaviors, user collaboration, and role analysis by enhancing the process model. From the c-level perspective, this category establishes organizational information to support identification aspects and answer questions (de Medeiros & van der Aalst, 2008) such as: how many employees are allocated in a specific process instance, what is the executors topology communication, how many transitions occur between roles, is there any communication hub, who delegates case, is there any vicious cycle of unnecessary work, and who execute the same role and tasks?

Important contributions have tried enriching the process model with role information with the ambition to partition activities while taking into account the originators and performers (Burattin, Sperduti, & Veluscek, 2013). This work also measures the handover of defined and employed roles by performing automatic partitioning extractions. Another relevant contribution established three types of organizational assistance focused on (1) Mining models of organizational structure, (2) Social network mining, and (3) Mapping information flows among allocated resources (machines or employees). These techniques were applied in Netherlands municipalities using the ProM framework (Song & van der Aalst, 2008).

#### *3.3.4. Concept drift*

Other more recent topics are related to detecting concept drift in processes that consists in: identifying the period when the change is occurring; localize the control-flow and/or resource involved and characterize (gradual or suddenly linear increasing) or; and, finally discovering the process change describing all in a perspective (Bose, van der Aalst, Zliobaite, & Pechenizkiy, 2014). Maaradji, Dumas, Rosa, and Ostovar (2017) explores two types of concept drift in a financial process: sudden and gradual drift.

#### *3.3.5. Decision mining*

Decision mining correlates context attributes to the process instances aiming to identify the usual rules in the gates. As example, we can observe a sales to cash process for a manufacturing, usually the decision gates to make a selling using the stock or using a production process, based on specifically (personalized) information of customers that usually asked for special products that is always not available on stocks.

According to Subramaniam et al. (2007), it is efficient consider as classes the outcomes at a decision point in a process model. It allows achieve significant precision in predictors for each gate or final process instance outcome state based on contextual data and historical event logs (Ghattas, Soffer, & Peleg, 2014).

#### *3.3.6. Resource allocation*

A previous systematic analysis identified many resource patterns used for process-aware information systems, such as Russell, van der Aalst, ter Hofstede, and Edmond (2005): direct, role-based, deferred, authorization, separation of duties, case handling, retain familiar, capability-based, history-based, organizational, automatic execution, distribution by Offer in a single or multiple resources, distribution by allocation, random, round robin, shortest queue, early distribution, distribution on enablement, late distribution, resource-initiated, resource-initiated execution, allocated work item, resource-initiated execution offered work item, system-determined work list management, resource-determined work list management, and selection autonomy.

Resource allocation is a topic explored in process mining for different goals, such as maximize the utilization, balance production, avoiding waste of resources, planning utilization and interventions (maintenances), among others. Many approaches combine machine-learning techniques to learn best decisions, e.g. Huang, Lu, and Duan (2011) focused in discovering predictive and useful resource allocation rules using association rule mining.

### *3.4. Supporting areas*

The supporting areas category group some crucial topics for use process mining in the real world. In order to implement process mining projects some topics are important, such as: understanding the main concepts (background), acquiring event data, following former methods for making process mining possible, using proposed architectures and tools, understanding possible ontologies and also knowing about lessons learned from other process mining applications already implemented. This category represents 33.82% of the papers analyzed, showing that process mining is not just a theoretical subject. Instead, many process mining projects have been conducted on different research areas, as it will be detailed on Section 5.

#### *3.4.1. Process mining applications*

This category group papers which describe applications of process mining without any relevant new contribution to the process mining techniques, algorithms or tools. This category contains 142 papers which indicates the use of a process mining technique without any new contribution to the techniques, tools or algorithms. These 142 papers will be further explored on Section 5.

#### *3.4.2. Methods for process mining projects*

Distinct methods have been proposed to conduct process mining applications and to combine process mining techniques with other methods. A well know paper published by van der Aalst (2011) proposed the L* life-cycle model consisting of five phases for conducting a process mining application.

- Plan and justify phase-focus is data-driven to explore and answer curiosities and gain insights; or, question-driven focused on explaining a specific situation e.g., why determined issues occur; or, goal-driven focused on improving the KPI, response time performance, explain deviations, reduce costs.
- Extract-explore the knowledge system data domain for recovering relevant events. Sometimes it can be very time-consuming owing to system complexity, e.g., with thousands of repositories spread among many tables.
- Discover a process model based on event logs-process miner algorithms are used to obtain a process model. When an understandable model is acquired, it is possible to start conformance checks, analyzing activities and deviations to the discovered process model.
- Create integrated process model-enhancements and added information are integrated in the model, promoting new perspectives related to time replayed, organization, resource utilization, case comparisons, and simulation, among other possibilities.
- Operational support-involves detecting, predicting, and recommending. This is the most advanced level of computational support, for example, the process mining tool should be capable to alert (email) on deviation cases, provide advice about bottlenecks, recommend resource setup or reallocation, etc.

Aimed to extend the traditional BP lifecycle, Delgado, Weber, Ruiz, de Guzmán, and Piattini (2014) proposed the BP Execution Measurement Model (BPEMM) for supporting continuous improvement. Their work analyzes and provides insight from a practical industry application. Here, process mining provide techniques that are very helpful. The proposed model is not limited to just obtaining insight, it can also help in many BP phases, e.g., performance evaluation.

Many relevant method proposals focused on specific application domains, for example: Mundbrod, Beuter, and Reichert (2015) focused on a method for systematic support of knowledge-intensive business processes involving more effective collaboration and coordination among employees. It was validated using development projects for electrical and electronic components. Mans, Reijers, Wismeijer, and van Genuchten (2013) applied a method that combines process mining with discrete-event simulation in healthcare to measure the effective of new technologies adoption. More precisely Mans (2013) was evaluating dentistry performance results for all chain, since patients to dental labs, according to the scenarios of traditional practices and usage of new technologies. Process mining was combined with discrete event simulation, allowing a faster start compared to the traditional simulation tools, reducing efforts and time to build simulation models.

#### 3.4.3. Architecture and tools

In this category, we aim to discuss published works for development of different tools to support and improve business processes. This involves many companies in development, moving significant amounts of money, large consulting services partners, and depends on complex, robust, and expensive infrastructures. Among these tools, we can cite BPMS, BPA, Business Intelligence (BI), Enterprise Performance Management (EPM), and BAM, among others. Since the 90s tools called WFMS have supported process execution and control. However, WFMS offers low support to identify process instances, measure the executed tasks, or to allow high-level process management. BPMS emerged as an evolution of WFMS to provide higher quality in process performance, diagnoses, and resources control (Ko, 2009). Typical BI tools are useful for monitoring process performance. These tools are usually data centered (data-centric), offering little or no process mining support, which is process centered (process-centric) (van der Aalst, 2016). Previous works such as the study by Guarda, Santos, Augusto, Silva, and Pinto (2013) considered integration between BI tools and process mining tools with the objective of delivering flexibility, quantitative analysis, statistic explanations, and to predict decision-making.

In the last decade, many tools have emerged to support process mining. However, the majority of publications use the ProM framework,[6] a very powerful and open source framework (van Dongen, de Medeiros, Verbeek, Weijters, & van der Aalst, 2005). ProM offers a pluggable architecture and is one reason why research communities embrace this framework. ProM allows flexibility in developing new algorithm (plug-ins), extending and combining them with standardized input and output formats. However, there are other options, such as Aris Process Performance Manager from Software AG,[7] BAB Framework[8] (Best Analytics of Big Data), Celonis Discovery,[9] Disco from Fluxicon,[10] Myivenio from Cognitive Technology,[11] Perceptive Process Mining from Lexmark (acquired by Hyland Software[12] and previously started by Pallas Athena as FLOWer)), Process Gold,[13] QPR ProcessAnalyzer, RapidProM,[14] SNP Business Process Analysis,[15] Signavio Process Intelligence,[16] UpFlux,[17] and others. Among this set of tools, it was observed that very few provide features for operational support as an active process control or as a native integration to BPMS providers.

In this category, we also identified the OLAP analytics engine to facilitate big process graph summarization, providing multiple views and dimensions at different granularity (Raichelson, Soffer, & Verbeek, 2017). Many works focused on improvements to the process miners horizontal scalability, combining Map-Reduce and Big Data engines such as Hadoop or Spark, among others. Other work combined stream processing for operational support (van Zelst, van Dongen, & van der Aalst, 2017).

Other research papers focused on definition of a standard format to abstract issues related to heterogeneity of the data sources. An approach to acquire, exchange, and analyze event logs was proposed, and this standard is called Extensible Event Stream (XES) (Xes, 2016). An XML Schema describing the structure and the extension standards is also available as a reference implementation in Java, called OpenXES.[18] A certification process was also developed to improve the portability among all commercial and academic process mining tools.

#### 3.4.4. Background

This category includes papers which cover general concepts of the process mining area, contributions to the area and secondary works.

#### 3.4.5. Gathering and cleaning

Gathering aims to generate of event logs from different data sources such as traditional information systems by observing focused data (Pérez-Castillo, Weber, de Guzmán, Piattini, & Pinggera, 2012a). The source for event acquisition can be a BPMS, traditional information system tables, system layers (e.g., DAO - Data Access Object), EDI documents, e-mail server, message hubs, API management services, web service infrastructures, Enterprise Application integration (EAI), message queues, among others. Other significant task for contribute with process mining results is cleaning event data. There many approaches for improving event logs quality since fix date-time using time window, fix activity labels with type mistakes, or any other imperfection patterns (Suriadi, Andrews, ter Hofstede, & Wynn, 2017). Despite the establishment of guiding principles to treat event data as First-Class Citizens (van der Aalst et al., 2012) there are common issues that are impacting the quality of process mining in real-life event logs (Bose, Mans, & van der Aalst, 2013). For instance, the events on the log can be at a different abstraction level of the activities on the process. Baier, Mendling, and Weske (2014) proposes an approach to deal with the abstraction level and improve the mapping between events and activities, which is a preprocessing step for improving the gathering of data.

#### 3.4.6. Ontology

Ontology and semantic-based approach can be combined to significantly improve the results achieved in different applications, such as discovery, conformance, trace clustering, task predictions, time prediction, decision mining, concept drift, among others. One generic approach was proposed in Kingsley, Tawil, Naeem, Islam, and Lamine (2016) aimed to enrich events logs with associated ontology structures, offering features to filter abstraction levels, and using query language.

## 4. Process discovery algorithms

The section provides a mapping of the algorithms used on the process discovery task. The first initiative to answer this question is categorized by the first published work for each algorithm. Fig. 8 presents an overview of process miners and lists some of the most relevant algorithms, including author and year.

The first relevant proposal was from Cook and Wolf (1995) describing three methods for process discovery based on finite state machines (FSMs). The three proposed process miners were based respectively, on a Recurrent Neural Network (RNet) from Das and Mozer (1994), the KTAIL algorithm from Biermann and Feldman (1972), and a statistical approach using Markov models to

[6] http://www.promtools.org.
[7] https://www.softwareag.com/.
[8] https://www.babcloud.org.
[9] https://www.celonis.com/.
[10] http://www.fluxicon.com/disco.
[11] https://www.my-invenio.com/.
[12] https://www.hyland.com/.
[13] http://processgold.com.
[14] http://www.rapidprom.org/.
[15] https://www.snp-ag.com/.
[16] https://processmining.signavio.com/.
[17] https://www.upflux.net.
[18] http://www.xes-standard.org/.

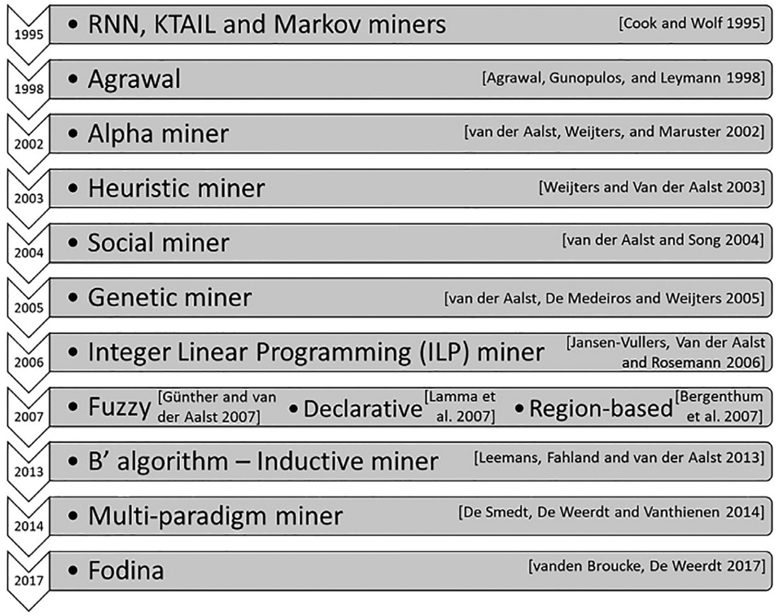


**Fig. 8.** Process mining algorithms by author and publication year.

identify the most probable events. In this paper, the authors already note the possibility to create models not only from event streaming processors but, also from any organizations process. They already discuss the continuous evolution of process models over time and how process discovery could help in their redesign. A second relevant work for discovering process models from event logs was published in 1998, aimed to generalize the problem of sequential patterns in a process (Agrawal et al., 1998). This proposal aimed to automatically generate a direct graph model of the process, delivering some quality measurements such as: completeness to preserve all the dependencies between activities that are in the log and allowing the execution of all process instances present in the log; non-redundancy to not introduce erroneous activity dependencies; and minimality to clarify the model presentation, building a minimal number of edges. The third most relevant work was by van der Aalst et al. (2002) where the Alpha Miner was proposed to find workflow models (WF-Net) and a tool called MiMo (Mining Module) based on ExSpect (EXecutable SPECification Tool) was introduced, supporting high-level Petri-nets. Mimo offers a graphical analyzer where the Alpha miner is used to automatically design the Petri net on-the-fly, including animations. The first proposal of the Alpha miner is not resilient to deal with noise, incomplete event logs, and was not capable of identifying short loops, mapping non-local dependencies, and deal with non-free choice constructs. Many researchers worked to improve and extend Alpha, proposing different algorithms to solve these limitations.

Other work that proposes a new approach involves process reverse engineering, in other words, instead of starting with process model design, it proposes starting by collecting event logs and then applying a new process mining technique for extracting the real executions flow (Weijters & van der Aalst, 2003). The proposal extends the Alpha algorithm and considers the frequency of the following relationship, calculating dependency/frequency tables to obtain a heuristic net, and the miner is called the Heuristic Miner. It is capable of handling noise, is frequency-based, and allows comparisons between manually designed model and executed process. This algorithm is the most used and customized, as it guarantees good fitness, but cannot provide total soundness because infrequent paths are not incorporated into the model.

A workflow mining system, called InWoLvE, was implemented by Herbst and Karagiannis (2004). The proposal was based on two steps: first, it produces a stochastic activity graph based on examples of observed process instances; then, as second step, it converts the graph to a process model using the ADONIS™ definition language, a registered trademark by BOC GmbH. The main algorithm proposed was splitPar, which optimizes the log-likelihood of samples by applying split-operations.

Another proposal focused on understating process deviation from a normal flow with different relationships patterns between individuals. In this sense, the sociometry and Social Network Analysis (SNA) from event logs of enterprise information systems were analyzed by van der Aalst, Reijers, and Song (2005), allowing analysis of social networks, identifying interaction patterns, and, evaluating the role of an individual in an organization. This work also proposed a new tool called MiSoN (Mining Social Networks) with integration to import logs from Staffware workflow™ and to three SNA tools.

The Genetic Process Mining approach was proposed by van der Aalst, de Medeiros et al. (2005) it is based on causal matrix as a representation for individuals and attempted to simulate the population evolution. It considers the relationship between Petri nets and the matrix representation to compute a genetic algorithms for discovering models from a given event logs. It applies elitism to select the fittest samples in the present generation that is preserved for the next generation, and uses the concepts of crossover and mutation to build the population elements.

Jansen-Vullers et al. (2006) created a new algorithm based on integer programming techniques. They show that it is possible to search for the best setting using an objective function, and applying integer programming techniques. The approach finds all solutions of a system of equations, and it is performed by integer programming techniques a minimizing function. To accomplish this, one step in advance is needed to transform the Event-driven Process Chain (EPC) into Event-driven Process Chains (C-EPCs). A two-step miner approach was proposed by Wil, van der Aalst, Rubin and Günther (2006), including in step one, a construction of a transition system, and, in the step two, the resulting transition system is transformed to a Petri net using the theory of regions. The step two was implemented by the ProM plug-in called Petrify to obtain a Petri net. Petrify is applied in many works for computing a Petri net and, it is possible extend to evaluate particular constraints.

A declarative miner (DecMiner) that combines Inductive Logic Programming (ILP) and Inductive Constraint Logic (ICL) algorithms was proposed by Lamma et al. (2007). ICL performs a learning function in which positive and negative interpretations are progressively transformed into rules, e.g., preconditions of activities. Each iteration of ICL discovers new clauses and transforms the rules that are added to the theory with their respective interpretations.

The second algorithm in number of papers is the Fuzzy miner, which offers an approach to simplification according to unary and binary significance. This approach is very useful for unstructured processes and for visualizing correlations of activities and connection measures (Günther & van der Aalst, 2007). This approach combines complex topology concepts from cartography, such as aggregation, abstraction, emphasis, and customization. However, the Fuzzy miner does not guarantee soundness nor fitness.

In the same year, process discovery based on regions was proposed. In region theory, a set of states are transformed into a transition system looking for agreement between all states on a determined region. In the state-based regions, sets of nodes, also called regions, are recognized that refer to a simplified section that can be used to construct a Petri net with the minimal net's behavior included (van Dongen, 2007). One of the proposed approaches from Bergenthum, Desel, Lorenz, and Mauser (2007a), was based on language-based regions and aimed to agree on language according to synthesis to resolve this process mining challenge.

An extensible framework was proposed by Leemans et al. (2013), called the B' algorithm and is known as an Inductive miner. The aim was discovering block-structured process models that are sound and fit to watched behavior on event logs. The algorithm characterizes the minimal information for discovering process model. The inductive miner provides a polynomial-time complexity delivering a feasible computational cost. The Inductive Miner appears to have high fidelity and is also considered suitable in the treatment of variability of event records for the abstraction of complex models. It also incorporates several criteria such as: frequency analysis, clustering, detection of deviations and frauds, analysis of times and bottlenecks, general vision, and process understanding and evaluation of values by outcome, all incorporated in the same solution (Leemans et al., 2013).

Proposed by De Smedt et al. (2014), Multi-paradigm miner is compounding by procedural and declarative constraints. This combination attempts to extract the best discovery from these two approaches, and after that, to provide insights into process models. This was performed mixing two previous algorithms, which are the Heuristics Miner and the Declare Miner.

Recently, vanden Broucke and Weerdt (2017), who proposed the Fodina algorithm, extended the most popular miner, the Heuristic algorithm, to include particular features. The approach achieves robustness to noisy, and also capable to identify duplicate activities. Another relevant contribution was flexibility, allowing to insert user choices for tuning the discovery process.

Fig. 9 shows the most used process mining algorithms. Note that the Heuristic miner and Fuzzy miner are the most used on applied cases. One possible explanation for this result is the capability of these algorithms to deal with noise and exceptions in unstructured processes (Günther & van der Aalst, 2007; Song et al., 2013). Among the hundreds of discovered process miner proposals, Fig. 9 presents the most popular approaches.

Among the most common tools, the ProM Framework has been cited in more than 402 papers, followed by Disco in 81 papers, and Inductive Workflow Learning via Examples (InWolvE) in 6 published papers. Other tools were also applied in some research, but are rare in published work. Some academic initiatives such as the ADEPT Workflow Management System, EMiT (Enhanced Mining Tool), InWoLvE, Little Thumb, MARBLE, MiMo (first supporting Alpha miner), and MiSoN (Mining Social Network) were also proposed over the years.

## 5. Process mining applications

This section is aimed at answering the question: Where is process mining being applied? The actual explored process mining application domains are very wide, we could observe that papers exclusively describing some process mining application represents almost 8% of the total number papers. Exploring all the analyzed papers we realized that the most relevant applications are in healthcare industry, hospitals, and clinical path. The second most relevant is information technology focused on applications in software development, maintenance, and other operation services. The manufacturing and industrial cases studied occupied the third position, followed by Education and Financial institutions. Section 5 describes the application domains and Fig. 10 presents the most active application domain according to number of publications.

Inside the 1278 target studies, we have classified 572 studies related to on application area and 12 studies related to more than one application area, and all of these studies have been full-text reading. For each one of the top six areas, we have empirically selected a group of studies to summarize the main contribution of process mining to the area, which is described by a table containing this grouping. Fig. 11 shows that the top six areas representing 79.41% of the total number of papers read.

- Healthcare: The application domain covering clinical path, patient treatment, or the primary processes of a hospital.
- ICT: Information and communication technology, related to software development, IT operation services, and telecommunication companies.
- Manufacturing: Application in industrial activities, realized by a factory that usually receives material and delivers partial or finished products. Among all industrial segments one very significant one is the automotive segment.
- Education: Application in Education, e-learning, scientific applications, and research centers with innovation process management.
- Financial: Application domain related to banks, insurance, in different processes such as payments, investments, deposits, risk analysis and mitigation, transfers, and other financial operations.
- Logistics: Application related to logistics, transportation, storage, and stock management.
- Public: Public administration, Government, Municipalities services, social security services, and postal services.
- Security: Security services in IT or related to process protocols, such as Maritime safety constraints.
- Call Center: Help desk and customer services for resolving reported issues or answering doubts.
- UX/Usability: Application to understand and improve the user experience focused in design, navigation path, websites and mobile application improvements, and e-Marketing.
- Robotics / Smart: New applications using advanced technologies related to scenarios, such as smart buildings, industry 4.0, and robotics.
- Entertainment: Applications for entertainment, sports, games, publicity, and news media.
- Utility: This segment involves energy generation or natural resources management, water treatment, balance, or delivery.
- Garment: Applications for clothes producers, such as dyeing and textile.

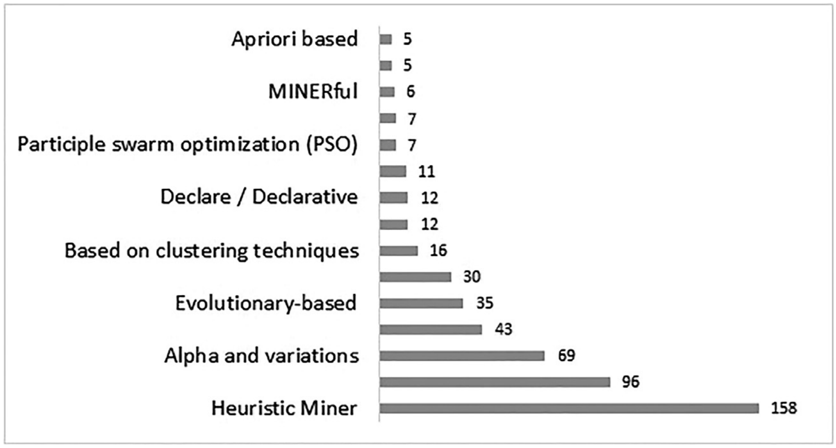


**Fig. 9.** Process mining algorithms distribution by use.

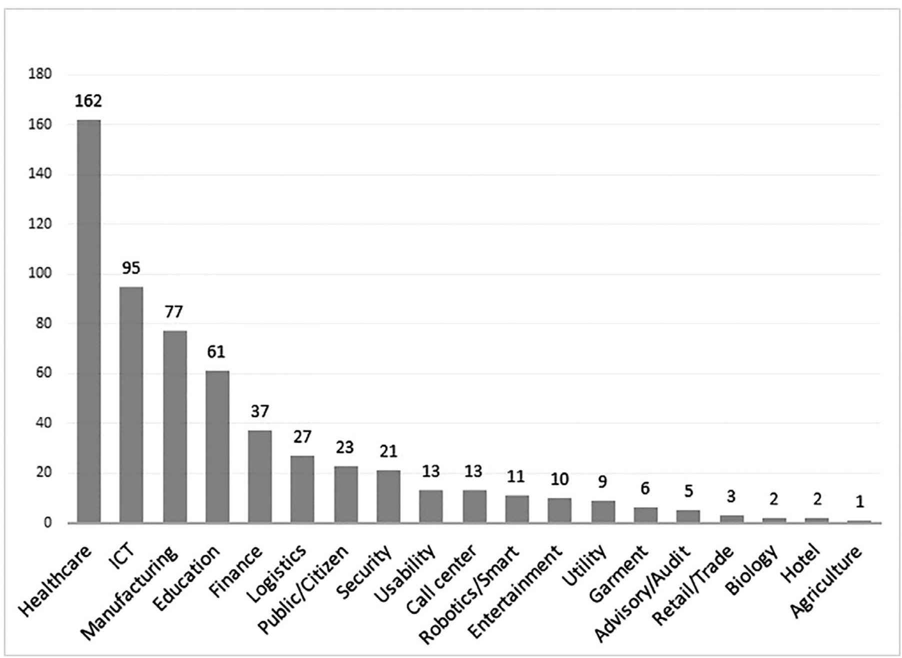


**Fig. 10.** Number of papers by application domain.

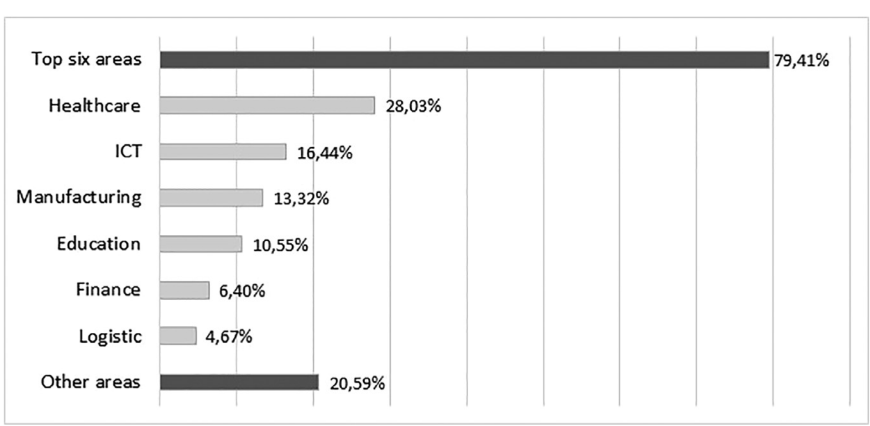


**Fig. 11.** Percentage of papers by application domain.

- Advisory: Consultancy services office related to advisory, audit, and taxes.
- Retail: Applications focused on offering and selling products.
- Biology: Processes in biological systems, such as metabolomics, biochemical networks having a set of metabolites, enzymes, reactions, and their interactions.
- Retail: Physical or virtual stores.
- Pharmacy / Biology: Related to biomedical applications, biological analysis, sequential assumptions
- Hotel: Segment focused on customer services, rooms, and hospitality.
- Agriculture: Agriculture and Food industry.

### 5.1. Healthcare

This section presents process mining applications focused on the healthcare segment, based on 162 articles selected, reporting relevant contributions and limitations in the development of solutions for this application domain. Many related studies indicate that the process models in healthcare differ from others in relation to their characteristics, high variability, complexity, security, privacy, multidisciplinary nature of their activities, and innovative and modern exams, treatments, or procedures, among others (Mans, van der Aalst, & Vanwersch, 2015; Munoz-Gama & Echizen, 2012; Rebuge & Ferreira, 2012). Process mining in healthcare is a relatively recent and its use can help in decision making, cost reduction, identification of patient flows for certain diseases, chances of treatment, maximization of flow with good results, correlations among administered treatments, quality of treatments, and complications, among others Rebuge and Ferreira (2012) and Rojas et al. (2016).

For the improvement of hospital care, the health area has been equipped with efficient information systems for clinical and administrative records of patient data. These collected data have been explored as an aid for managers and clinicians in decision making, aiming to provide more efficient and effective services (Mans et al., 2013; Mans et al., 2015). Process mining has been one of the techniques that provide most benefits from the availability of data recorded in hospital information systems (Ferdinand & Emmanuel, 2015; Huang, Dong, Bath, Ji, & Duan, 2014). Through the process mining application, hospital organizations can evaluate how clinical pathways are conducted, check whether certain clinical guidelines and medical protocols have actually been followed, visualize resource utilization, identify bottlenecks, obtain insight, and other aspects of performance (Cho, Song, & Yoo, 2014; Rovani, Maggi, de Leoni, & van der Aalst, 2015).

In the area of healthcare there is a fairly consolidated term, which is the clinical pathway. A clinical pathway can be considered as a specific process, according to the patient's illness, which describes the flow of attendance of this patient (Huang, Lu, Duan, & Fan, 2013). Numerous studies show the application of process mining as a useful tool to verify clinical pathways adopted in hospital settings (Baker et al., 2017; Caron et al., 2014a; Fernandez-Llatas, Martinez-Millana, Martinez-Romero, Benedi, & Traver, 2015; Fernandez-Llatas, Valdivieso, Traver, & Benedi, 2014; Huang, Dong, Ji et al., 2014).

The first research of a process mining application in healthcare selected by the literature review was in 2003, and presents the importance of resource analysis using process discovery and comparing it with clinical guidelines for conformance analysis (Ciccarese et al., 2003). However, the authors encountered difficulties regarding the representation of the discovered models because algorithms intended for the characteristics of the segment structured process as in industrial domain were applied, which do not fit the unstructured models typical of healthcare.

Different applications have focused on the discovery of processes for comparison with clinical guidelines for process improvement identification. One common obstacle is determining approaches capable of dealing with complexity and difficulties regarding the issues involved with generating complex process models and high levels of variability, also termed spaghetti models.

The description of roles, and iteration and collaboration among hospital emergency department professionals is a very common application in which the Fuzzy Miner algorithm (Günther & van der Aalst, 2007) constitutes an important approach to deal with the inherent unstructured characteristics of the discovery of healthcare process models. The adopted strategy provided visualization through the abstraction of reality and the representation of the most relevant activities, and associated the temporal measures between them. A preprocessing technique emerged to establish ways of clustering activities to represent complex models, such as the MinAdept algorithm (Li et al., 2010), which aims to make the process clear to compare process variants in healthcare. The process variant are groups of equals or similar paths that considers behavioral or structural similarity to split in simpler models. It is then much easier to realize conformance analysis from the discovered process models to the reference model.

The high variability evidenced in healthcare research is directly related to observation regarding patient pathways, where it is difficult to establish any predictability of their script, as it is determined by many causes, such as; biological interactions, pathology, and treatments carried out, among others. Besides this lack of predictability, pathways are stochastic, that is, random in relation to their occurrence and are difficult to plan for Prodel, Augusto, Xie, Jouaneton, and Lamarsalle (2015).

Presented techniques often deal with questions of analysis directed at a specific context, pathology, or surgical procedure, which require time for the definition of parameters and acquired knowledge by healthcare specialists. Besides the different applications of the different techniques, it important to note that most research studies were carried out in a single institution, usually in a specific hospital, so that there is no visualization of the global processes of patient care pathways, making it difficult for quality healthcare assessment including measurement outcomes of any patient.

Another area of health that has benefited from the application of process mining is so-called emergency medicine (Alvarez et al., 2018; Basole, Braunstein et al., 2015; Kim et al., 2013; Rojas et al., 2017). This area is dedicated to the diagnosis and treatment of unexpected diseases and accidental injuries, embracing the initial evaluation of the patient to verify the complexity of their health, diagnosis, treatment, and coordination of care among various providers at the disposal of patient in need of surgical or non-surgical medical care (Alvarez et al., 2018).

The Janus approach Cecconi, Ciccio, Giacomo, and Mendling (2018) presents the reactive restrictions discovery using an online mode. The reactive restrictions are activated by reading an historic event log and they make predictions based on the historical information. The solution is shown efficient evaluating the restrictions in comparison to the event log and the normative model. It is an interesting approach to generate online alerts to the health attention area.

Process mining applications in healthcare represent a potential benefit, allowing specialists, through a patient's event logs, to manage the process model, evaluate compliance with clinical guidelines and protocols, and identify points of process improvements. There are also many opportunities for future work in prediction and recommendation (Yang, Dong et al., 2017). Aside from the effort to apply techniques to reduce variability and establish a more regular protocol (reference model), it is important to note that sometimes patients are saved by taking shortcuts as physicians deviation from standard procedures, but protocol deviations sometimes result in fatalities. Certainly, data analysis and process mining tools can support improvements in comprehension, performance results and effectiveness (van der Aalst, 2016).

Table 3 provide an overview of the process mining studies applied to the healthcare domain.

### 5.2. Information and communication technology (ICT)

From process mining applications 95 papers focused on the IT segment, focused on different goals such as process improvement, process evaluation, agile task management, collaboration, increased performance maintenance, and compliance checks in operation and service management, among others.

Incident and change management processes were analyzed based on supporting IT system events in a case study at a IT outsourcing organization regarding compliance with an IT Infrastructure Library (ITIL) (Baier et al., 2014). They used domain knowledge gathered in documentation using a semi-automatic activity mapping, following four steps: annotation in process activities; mapping of level that employs a domain expert to define sub-activity levels and events; outlining of related data context; and, finally clustering of events for findings. One relevant study focused on a more unstructured process for software defects and issued resolutions by conducting a case-study for the open-source Firefox browser (Juneja, Kundra, & Sureka, 2016). The authors proposed a trace clustering approach for reducing structural complexity and improving the process models comprehensibility by process analysts. All event logs were extracted from the Bugzilla issue tracking system after a sequential analysis was performed using an adaptation of the K-medoid algorithm. The used distance function were Longest Common Subsequence (LCS) and Dynamic Time Warping (DTW) to obtain clusters with good intra-class similarity. Then, each cluster was analyzed using a fuzzy miner algorithm available in the Disco tool, and, finally fitness and complexity metrics compared the goodness of each process model. As a result, the authors could successfully show that clustering enables easy identification of bottlenecks and checks loops in reopening issues, among other analyses.

Šamalíková, Trienekens, Kusters, and Weijters (2009), Samalikova (2012), Rubin et al. (2007) and Samalikova, Kusters, Trienekens, and Weijters (2014) point out the promising that process mining can achieve when applied in software process evaluations considering current limitations such as complexity and instability. These works use records of information systems for analysis and evaluation of software processes. The resulting analysis opportunities are based on considering the CMMI guidelines (Capability Maturity Model Integration) for evaluation. Aside from this perspective Valle, Santos, and Loures (2017) used process mining to support the evaluation of Standard CMMI Appraisal Method for Process Improvement (SCAMPI).

Another relevant application with more focus on quantitative analysis explored a database with event logs from five years of software development process execution, and 2000 process cases in a large Brazilian company (Lemos, Sabino, Lima, & Oliveira, 2011). This study showed many evidences that process mining is capable to improve software process maturity level. The applied process mining tool uses a Markov Chain, implemented on ProM, for sequence clustering analysis. The authors analyzed the opportunity for using process mining to the evaluation of the Quantitative Project Management (QPM) process area, and found that it can effectively establish an infrastructure for Statistical Process Control (SPC) for higher levels of maturity.

Rubin, Lomazova, and van der Aalst (2014) applied process mining for analysis of software processes and as a stage in their life cycle. The goal is to provide an agile software development methodology. Gupta (2014) proposed applying process mining to derive execution-time process models, to identify and remove deviations, to expand the resources of current software engineering tools, and making stakeholders more aware of the process, to understand the pattern of interaction, and to improve the efficiency of a software development. The same process mining framework, named Nirikshan, was evaluated using bug report history data (Gupta & Sureka, 2014a). Gupta (2017) also proposed the improvement of the software maintenance process through the analysis of records using process mining.

Table 4 provide an overview of the process mining studies applied to the ICT domain.

### 5.3. Manufacturing

In this industrial segment, 77 research papers were selected. Owing to the increasing presence of information systems, such as enterprise resource planning (ERP), manufacturing execution systems (MES), supervisory control and data acquisition (SCADA), and operations and logistics, various studies have explored process mining using events in this area. One of the first research studies in this area was reported by Ho and Lau (2007), and focused on improving a discrete production process to discover hidden relationships and prove suggestions from a vast amount of data events. The proposed methods offered higher flexibility on production process management with decision support capability and improvement in a real-time manner. The work of Rozinat, de Jong et al. (2009) also applied process mining related to industrial equipment rather than to administrative processes. In this paper, the authors analyzed the event records of a manufacturer of chip manufacturing equipment, and the challenge was to analyze a less-structured process when compared to administrative processes.

The work of Li, Liu, and Fan (2008) and Li, Liu, Yin, and Zhu (2010) analyzed the process of knowledge management in industrial maintenance. They also analyzed eligible resources capable to carry out same types of tasks, identifying organization clusters, and detecting the most relevant relationships. Karray et al. (2014) proposed a system called PETRA for the analysis of maintenance processes. The idea is to use process mining as a behavior verification tool and to extract rules related to the execution of maintenance activities. This work combines IMANO ontology for organizing actions into 12 types of maintenance, named as: corrective, preventive, scheduled, pre-determined, condition-based, predictive, remote, deferred, immediate, on-line, on-site, and operator (Karray et al., 2014).

A recent application was conducted by Ruschel, Santos, and de Freitas Rocha Loures (2017), which integrates process mining with a Bayesian network of predictive models for determining maintenance intervals for industrial equipment. The idea is to provide a tool to support managers in decision making regarding maintenance equipment stops, reducing downtime, and avoiding unplanned stops.

Table 5 provide an overview of the process mining studies applied to the manufacturing domain.

### 5.4. Education

Many process mining applications are aimed to discover, monitor, and improve education processes, representing 61 papers. During the learning process, it is possible to analyze the set of recurrent behaviors that can be found in the events. Some papers focus on looking for appropriate learning paths based on a user profile group, detecting learning styles and desired content type, identifying trends in online learning studies, recommending a learning path for best results, analyzing relationships between students, supporting instructors to increase their awareness toward students'

**Table 3**
Healthcare studies.

| Main Contribution | Studies |
|---|---|
| Conformance checking evaluation based on clinical protocols and guidelines | Alizadeh, Lu, Fahland, Zannone, and van der Aalst (2018), Giacalone, Cusatelli, and Santarcangelo (2018), Funkner, Yakovlev, and Kovalchuk (2017a), Kukreja and Batra (2017), del Pilar Villamil et al. (2017), Prodel, Augusto, Xie, Jouaneton, and Lamarsalle (2017), Riaño and Ortega (2017), Yang, Zhou et al. (2017), Zhou et al. (2017), Chomyat and Premchaiswadi (2016), Erdogan and Tarhan (2016), Lismont et al. (2016), Xu, Jin, and Wang (2016), Basole, Park et al. (2015), de Leoni, Maggi, and van der Aalst (2015), Ferdinand and Emmanuel (2015), Huang, Dong, Ji, Yin, and Duan (2015), Letia and Goron (2015), Rovani et al. (2015), Zeng, Lu, Liu, Duan, and Zhou (2015), Caron et al. (2014a), Defossez, Rollet, Dameron, and Ingrand (2014), Kelleher, Bose, Waterhouse, Carter, and Burd (2014), Montani, Leonardi, Quaglini, Cavallini, and Micieli (2014), Liu, Liu, Li, Xie, and Lakshmanan (2014), qiong Wang, shu Zhou, li Tian, ming Qian, and song Li (2014), Kelleher, Bose, Waterhouse, Carter, and Burd (2013), Peleg (2013), Sánchez-Garzón, González-Ferrer, and Fernández-Olivares (2013), Huang, Lu, and Duan (2012c), Rebuge and Ferreira (2012), Li et al. (2010), Mulyar, van der Aalst, and Peleg (2007), Ciccarese et al. (2003) |
| Customer satisfaction assessment | van Genuchten, Mans, Reijers, and Wismeijer (2014), Kim and Lee (2013) |
| Method for healthcare analysis using process mining | Cho, Song, Comuzzi, and Yoo (2017), Mans et al. (2013), Huang, Zhu, and Wu (2012) |
| Outliers Detection | Wolf, Herrmann, and Rothermel (2013), Bouarfa and Dankelman (2012), Han, Jiang, and Cai (2011) |
| Performance and bottleneck analysis, and time/scheduling management | Abohamad, Ramy, and Arisha (2017), Ganesha, Soundarya, and Supriya (2017), Jangvaha, Porouhan, Palangsantikul, and Premchaiswadi (2017), Jaturogpattana, Arpasat, Kungcharoen, Intarasema, and Premchaiswadi (2017), Montani, Leonardi, Striani, Quaglini, and Cavallini (2017), Saelim, Porouhan, and Premchaiswadi (2016), Toyawanit and Premchaiswadi (2016), Tsumoto and Hirano (2016), Yampaka and Chongstitvatana (2016), Yoo, Jung et al. (2016), Yoo, Cho et al. (2016), Forsberg, Rosipko, and Sunshine (2015), Jaroenphol, Porouhan, and Premchaiswadi (2015), Tsumoto and Hirano (2015), Rebuge, Lapao, Freitas, and Cruz-Correia (2013), Mans, Reijers, van Genuchten, and Wismeijer (2012), Pla, Gay, Meléndez, and López (2012), Blum, Padoy, Feußner, and Navab (2008), Maruster and Jorna (2005) |
| Post-processing (dealing with healthcare complex process models) | Mapikou and Etoundi (2016), Ferreira et al. (2007) |
| Predictive analysis for managing cases and diseases | Lakshmanan, Mukhi, Khalaf, Martens, and Rozsnyai (2012), Fernandez-Llatas et al. (2011), Francescomarino, Dumas, Maggi, and Teinemaa (2017), Huang, Dong, Ji, He, and Duan (2016), Zhang and Chen (2012), Siddiqui et al. (2015), van der Spoel, van Keulen, and Amrit (2013), Dagliati et al. (2014) |
| Pre-processing (data gathering, dimensionality reduction, data quality, OLAP) | Rojas et al. (2017), Suriadi et al. (2017), Rogge-Solti, Mans, van der Aalst, and Weske (2013), Song et al. (2013), Vogelgesang and Appelrath (2013), Zhang, Martikainen, Pulli, and Naumov (2011), Su and Al-Hakim (2010) |
| Process discovery and analysis (clinical protocol and guideline) for care evaluation | Chen et al. (2018), Prodel et al. (2018), Baker et al. (2017), Chen, Yang, Zhou, Burd, and Marsic (2017), Dagliati et al. (2017), Funkner, Yakovlev, and Kovalchuk (2017b), Gatta et al. (2017), Halioui, Martin, Valtchev, and Diallo (2017), Ma'arif (2017), Metsker, Bolgova, Yakovlev, Funkner, and Kovalchuk (2017), Toth, Machalik, Fogarassy, and Vathy-Fogarassy (2017), Wang, Tian, Yu, Qi, and Yang (2017), Xu, Jin, Wei, and Wang (2017), Yang, Dong et al. (2017), Augusto, Xie, Prodel, Jouaneton, and Lamarsalle (2016), Fernandez-Llatas, Bayo, Martinez-Romero, Benedi, and Traver (2016), Garg and Agarwal (2016), Lu, Zeng, and Duan (2016), Meng, Ooi, Soh, Teow, and Kannapiran (2016), Vitali and Pernici (2016), Xu, Jin, Wei, Lv, and Wang (2016), Basole, Braunstein et al. (2015), Benner-Wickner, Brückmann, Gruhn, and Book (2015), Delias, Doumpos, Grigoroudis, Manolitzas, and Matsatsinis (2015), Fernandez-Llatas et al. (2015), Garcia et al. (2015). Meier, Dietz, Boehm, and Neumuth (2015), Neamsirorat and Premchaiswadi (2015), Papadopoulos et al. (2015), Prodel et al. (2015), Yao et al. (2015), Zhang, Padman, and Patel (2015), Caron et al. (2014b), Cho et al. (2014), Fernandez-Llatas, Sacchi et al. (2014), Fernandez-Llatas, Valdivieso et al. (2014) Huang, Dong, Ji et al. (2014), Huang, Dong, Bath et al. (2014), Huang, Dong, Duan, and Li (2014), Klausner et al. (2014), Perimal-Lewis, De Vries, and Thompson (2014), Zhou, Wang, and Li (2014), Caron, Vanthienen, and Baesens (2013b), Huang, Lu, and Duan (2013), Huang, Lu, Duan, Fan (2013), Kim et al. (2013), Perimal-Lewis, Qin, Thompson, and Hakendorf (2013), Comuzzi, Vonk, and Grefen (2012), Fei and Meskens (2012), Kaymak, Mans, van de Steeg, and Dierks (2012), Perimal-Lewis, Qin, Thompson, and Hakendorf (2012), Fernańdez-Llatas, Meneu, Benedi, and Traver (2010), Micieli, Cavallini, Quaglini, Fontana, and Duè (2010), Zhou and Piramuthu (2010), Tesanovic, Manev, Pechenizkiy, and Vasilyeva (2009), Bergenthum, Desel, Lorenz, and Mauser (2007b), Li, Yuan, and Kong (2007), Xing et al. (2007), Curia, Gallucci, and Ruffolo (2005), Aalst (2001) |
| Proposal for process discovery algorithm | Cecconi et al. (2018), Chabrol, Dalmas, Norre, and Rodier (2016), Cheng, Ou-Yang, and Juan (2012), Pérez-Castillo, Weber, de Guzmán, Piattini, and Pinggera (2012b) |
| Resource or team allocation | Alvarez et al. (2018), Ganesha, Dhanush, and Raj (2017), Ganesha, Supriya, and Soundarya (2017), Ganesha, Raj, and Dhanush (2017), Jaisook and Premchaiswadi (2015), Krutanard, Porouhan, and Premchaiswadi (2015), Meethaisong and Premchaiswadi (2015), Partington, Wynn, Suriadi, Ouyang, and Karnon (2015), Rattanavayakorn and Premchaiswadi (2015), Zhao, Liu, Dai, and Ma (2015), Guo, Brown, and Rasmussen (2013) Zeng, Sun, Duan, Liu, and Wang (2013), Huang et al. (2011), Darabi, Galanter, Lin, Buy, and Sampath (2009) |
| Secondary studies or studies describing process mining concepts for healthcare | Ghasemi and Amyot (2016), Kurniati et al. (2016), Rojas et al. (2016), Mans et al. (2015), Yang and Su (2014), Homayounfar (2012) |

**Table 4**
Information and communication technology studies.

| Main Contribution | Studies |
|---|---|
| Application of Map Miner Method, generating an intentional process map, to identify developers' behaviors during the development process | Khodabandelou, Hug, and Salinesi (2014), Khodabandelou, Hug, Deneckère, and Salinesi (2014) |
| Application of process discovery on an ICT process context | Acampora, Bernardi, Cimitile, Tortora, and Vitiello (2017), Bobek and Nalepa (2017), Chindenga, Scott, and Gurajena (2017), Du, Cai, Jiang, and Huang (2017), Mitsyuk (2017), Rosa (2017), Rosa, Campos, and Cavalcanti (2017), Zerouali and Mens (2017), Blum (2016), Joe, Emmatty, Ballal, and Kulkarni (2016), Lopez, Maag, Saint-Pierre, Bustos, and Cavalli (2015), Shah, Khadke, and Rana (2015b), Atastina and Kurniati (2014), Despaux, Song, and Lahmadi (2014), Joishi and Sureka (2014), Mastroianni and Papuzzo (2014), Saint-Pierre, Cifuentes, and Bustos-Jimenez (2014), Wang, Jiang, and Cai (2014), Fahland, Lo, and Maoz (2013), Javed, Abgaz, and Pahl (2013), Xu, Xu, Bavikati, and Wong (2012), Goedertier, Weerdt, Martens, Vanthienen, and Baesens (2011), Haung, Chen, and Chung (2011), Poncin, Serebrenik, and van den Brand (2011b), Edgington, Raghu, and Vinze (2010), Lou, Fu, Yang, Li, and Wu (2010), Alonso, Fuente, and Brugos (2009), Gillblad, Steinert, and Ferreira (2009), Kanstren (2009), Misev and Atanassov (2009), Sengupta, Banerjee, Bisdikian, and Hurley (2008) Shao, Chen, Tao, Yan, and Anerousis (2008), Gönczy, Heckel, and Varró (2007), Ogasawara, Tayama, and Yamamura (2006) |
| Application of process discovery on an ICT process to improve process development and/or software maintenance | Damevski, Shepherd, Schneider, and Pollock (2017), Gupta (2017), Lübke (2017), Damevski, Chen, Shepherd, and Pollock (2016), Weber, Nepal, and Zhu (2016), Blum, Simmonds, and Bastarrica (2015), Santos, Oliveira, and e Abreu (2015), Shah, Khadke, and Rana (2015a), Gupta (2014), Gupta and Sureka (2014a), Gupta, Sureka, and Padmanabhuni (2014), Mittal and Sureka (2014), Chen, Hoi, and Xiao (2011), Duan and Shen (2011), Peŕez-Castillo, Weber, de Guzmań, and Piattini (2011), da Costa Cordeiro et al. (2009), Zhang and Hochstein (2009) |
| Application of process discovery to analyse mobility patterns of social media users | Diamantini, Genga, Marozzo, Potena, and Trunfio (2017) |
| Application of process discovery to map user-system interactions for improving software design and development and/or evaluate system's performance | Dabrowski et al. (2017), Rubin et al. (2014) |
| Application of process mining techniques to ICT process discovery for software process assessment, improvement, and conformance checking (against the desired model) | Gupta, Asadullah, Padmanabhuni, and Serebrenik (2017), Valle et al. (2017), Keertipati, Licorish, and Savarimuthu (2016), Mahendrawathi, Astuti, and Nastiti (2015), Samalikova et al. (2014), Sebu and Ciocarlie (2014), Wang, ter Hofstede et al. (2014), Xu, Zhu, Weber, Bass, and Sun (2014), Sunindyo and Ekaputra (2013), Astromskis, Janes, and Mahdiraji (2012), Samalikova (2012), Lemos et al. (2011), Poncin, Serebrenik, and van den Brand (2011a), Samalikova, Kusters, Trienekens, Weijters, and Siemons (2010), Sunindyo, Moser, Winkler, and Biffl (2010a), Akman and Demirörs (2009), He, Guo, Wang, and Guo (2009), Šamalíková et al. (2009), Rubin et al. (2007), Huo, Zhang, and Jeffery (2006b), Huo, Zhang, and Jeffery (2006a), Colombo, Damiani, and Gianini (2006), Cook and Wolf (1998) |
| Application of process mining techniques to improve the security of web information systems | Bernardi, Alastuey, and Trillo-Lado (2017), Compagna, dos Santos, Ponta, and Ranise (2017) |
| Clustering events for improving the discovery of ICT processes | Jlailaty, Grigori, and Belhajjame (2017), Juneja et al. (2016), Robinson and Deng (2015), Baier et al. (2014) |
| Cross-Organizational Process Mining Framework to compare organizations based on the usage of a software product (ERP) | Aksu, Schunselaar, and Reijers (2016) |
| Detect deviations from ICT processes in real time | Sebu and Ciocarlie (2015) Xu et al. (2013) |
| Discover behavioral model per component for software execution process | Liu, van Dongen, Assy, and van der Aalst (2016) |
| Discover collaboration patterns on software development processes | Yu and Wang (2017), Fan, Li, and Zhao (2012) |
| Event log enhancement approach based on human work patterns to improve quality of the discovered models | Awad, Zaki, and Francescomarino (2016) |
| Predict problems related to performance or conformance of running software based on process mining techniques | Aalst (2015) |
| Process cube modelling for defect resolution process, allowing the performance of a set of OLAP operations | Gupta and Sureka (2014b) |
| Recommendation system based on clustering techniques and discovered processes | Hug, Deneckere, and Salinesi (2012) |

collaboration, for diagnosing bottlenecks, analyzing performance, checking conformance, and monitoring students, among others.

Trcka, Pechenizkiy, and van der Aalst (2011) describes how process mining techniques can help discovering or analyzing the complete educational process, by extracting knowledge from educational information systems to identify students behavior. They focused on process discovery and conformance checking using the ProM tool in a simplified context containing examination traces.

An approach by Okoye, Tawil, Naeem, Bashroush, and Lamine (2014) and Okoye, Tawil, Naeem, Bashroush, and Lamine (2014) focused on discovering patterns and rules through semantic reasoning based on Web Ontology Language (OWL) and Semantic Web Rule Language (SWRL) with Protègè. This work applied process metrics and design strategies based on semantic rule-based to provide a useful supervised guide considering individual differences for learning. The rules and process mining techniques were combined for behavioral detection in a continuous process to support and anticipate pattern changes according the user. The results suggest a promising learning support, capable to deliver filtered, fitted and personalized content according user behavior.

**Table 5**
Manufacturing studies.

| Main Contribution | Studies |
|---|---|
| Approach of data cleaning to separate the multi-version process information in real-life data set for process discovery | Zhaoxia, Jianmin, Lijie, and Yingbo (2009) |
| Conformance checking applied on manufacturing processes | Hsu, Chuang, Lo, and He (2017), Haung, Chen, and Chung (2013), xia Wang, min Wang, chen Zhu, and jie Wen (2012), Zha, van der Aalst, Wang, Wen, and Sun (2011), van der Aalst (2006) |
| Decision support for resource distributions based on process mining techniques | Liu, Yi, Ni, and Liu (2008) |
| Discriminative rule mining to discover process containing non-atomic activities | Bernardi, Cimitile, Francescomarino, and Maggi (2016) |
| End to end delay computation based on Markov chain | Despaux, Song, and Lahmadi (2015) |
| Evaluating different aspects of manufacturing process flows and address practical challenges of state-of-the-art industrial process mining | Flath and Stein (2018), Li, Chan, Liang, and Luo (2015), kyung Lee et al. (2013), Huang, Li, Yin, and Zhao (2012), Ingvaldsen and Gulla (2012), Karnok and Monostori (2011), Viale, Frydman, and Pinaton (2011), Ou-Yang and Juan (2010) |
| Event-based predictions for manufacturing processes | Krumeich, Jacobi, Werth, and Loos (2014) |
| From databases of production information systems, process mining is used to provide guidelines for the improvement of underperforming BPMs in relation to the manufacturing processes | Park, Lee, and Zhu (2014), Yang, Park, Cho, Song, and Kim (2014) |
| Novel process mining anomaly detection method for identifying anomalous behaviour and cyber-attacks using ICS data logs and the conformance checking analysis technique from the process mining discipline | Myers, Suriadi, Radke, and Foo (2018) |
| Organizational perspective analysis on manufacturing process | Jin, Wang, and Wen (2007) |
| Optimizing spaghetti process models | Chinces and Salomie (2015) |
| Process discovery applied on manufacturing context | Keski-Valkama (2017), dos Santos, Piechnicki, de Freitas Rocha Loures, and Santos (2017), Bruntsch and Tseng (2016), Ceravolo et al. (2016), da Silva, Pereira, and Götz (2016), Dohrmann (2014), Engel and Bose (2014), Ghattas et al. (2014), Krathu et al. (2014), Estrada-Vargas, Lopez-Mellado, and Lesage (2013), Yano, Nomura, and Kanai (2013), Kondo, de F. R. Loures, and Santos (2012), Schwenke, Wagner, Gellrich, and Kabitzsch (2012), Viale, Giambiasi, Frydman, and Pinaton (2012), Wagner, Schwenke, and Kabitzsch (2012), Wang, Tan, Wen, Wong, and Guo (2012), Li, Reichert, and Wombacher (2011), Worgan, Behera, Cohn, and Hogg (2011), Sunindyo, Moser, Winkler, and Biffl (2010b), Berlingerio, Pinelli, Nanni, and Giannotti (2009) Lau, Ho, Zhao, and Chung (2009b), Fu and Peng (2008), Tupa (2008), Yi and Lai (2008) |
| Process mining is focused in discover problems related to green factory and lean manufacturing | Markowski and Przybyłek (2017), Jo, Noh, and Cho (2014) |
| Process mining is used to analyze and improve the maintenance management of industrial systems | Ruschel, Santos, and de Freitas Rocha Loures (2018), Ruschel et al. (2017), Karray et al. (2014), Li et al. (2010), Li et al. (2008) |
| Process mining is used to analyze knowledge workers, operator/users actions and users' knowledge | Gadler, Mairegger, Janes, and Russo (2017), Hu, Al-Dabbagh, Chen, and Shah (2016), Yahya, Song, Bae, ook Sul, and Wu (2016), Mundbrod et al. (2015) |
| Process mining is used to analyze the design of industrial products and to determine products which usually cause delays at production machines | Bartík and Pospíšil (2015), Berriche, Zeddini, Kadima, and Riviere (2015) |
| Process mining is used to quality enhancement in manufacturing systems | Dasani, Shah, Chen, Funnell, and Pollard (2015), Lau, Ho, Chu, Ho, and Lee (2009a), Ho and Lau (2007), Ho, Lau, Lee, Ip, and Pun (2005) |
| Proposal of a method to analyze and predict cost of manufacturing process | Tu and Song (2016) |
| Recovering business processes from systems based on process mining techniques | Pérez-Castillo, Cruz-Lemus, de Guzmán, and Piattini (2012), Sailer, Deubzer, Luttgen, and Mottok (2016) |
| Similarity measure between process models on manufacturing context | Zha, Wang, Wen, and Wang (2009) |
| Process discovery applied to simulate and optimize process model on manufacturing context | Wu, Lai, and Sun (2006) |
| Recommendation of process mining algorithms | Wang, Wong, Ding, Guo, and Wen (2012) |
| Recommendation support for product-based workflow | Vanderfeesten, Reijers, and van der Aalst (2011) |
| Using a variety of process mining techniques, three different perspectives are analyzed: the process perspective, the organizational perspective, and the case perspective | Meincheim, dos Santos Garcia, Nievola, and Scalabrin (2017), R'bigui and Cho (2017), Rozinat, de Jong et al. (2009), van der Aalst et al. (2007) |

Groba, Barreiros, Lama, Gewerc, and Mucientes (2014) conduct an applied research study presented a learning approach for personal learning environments based on a social network aimed to enable teachers to intuitively analyze student flows using a graphical interface. The developed tool, named SoftLearn, was applied at the University of Santiago de Compostela, Spain, and it allowed contextualizing and following students individually and in groups, making it easier for improvements in the overall learning process, recommending learning material, fixing sequences, and filling in attendance slots. Barreiros, Lama, Mucientes, and Vidal (2014) described the SoftLearn tool in more detail.

Table 6 provide an overview of the process mining studies applied to the educational domain.

**Table 6**
Education studies.

| Main Contribution | Studies |
| --- | --- |
| Automatic behavioral feedback for learning based on process discovery | Sedrakyan and Snoeck (2017), Calvo, O'Rourke, Jones, Yacef, and Reimann (2011) |
| Conformance checking and performance analysis on educational support processes | Anuwatvisit, Tungkasthan, and Premchaiswadi (2012) |
| Clustering to improve educational process discovery (students profile mapping) | Ariouat, Cairns, Barkaoui, Akoka, and Khelifa (2016), Cairns, Gueni, Hafdi, Joubert, and Khelifa (2015), Bogarín, Romero, Cerezo, and Sánchez-Santillán (2014) |
| Discover of semantic association rules for the students behavior within learning process (including self-regulated learning processes) | Okoye et al. (2014), Okoye et al. (2014) |
| Identification of association rules to discover students learning patterns | Huang, Lien, and Wang (2016) |
| Integration of process mining and simulation techniques in a process redesign project | Aguirre, Parra, and Alvarado (2013) |
| Process discovery applied to educational support processes | Ayutaya, Palungsuntikul, and Premchaiswadi (2012), Potavin, Jongswat, and Premchaiswadi (2012), Weerapong, Porouhan, and Premchaiswadi (2012) |
| Process discovery to analyse and support innovation processes | Diamantini, Genga, Potena, and Storti (2013), Genga (2013) |
| Process discovery to identify learning processes and improve them by a remodeling algorithm | Li (2009) |
| Process discovery to visualize group interactivity for collaborative learning processes (including self-regulated learning) | Sedrakyan, Malmberg, Verbert, Järvelä, and Kirschner (2018), Porouhan and Premchaiswadi (2016), Schoor and Bannert (2012) |
| Process mining to discover and check conformance of the students behavior within a learning process - Curriculum | Trcka et al. (2011), Priyambada, Mahendrawathi, and Yahya (2017), Azzini, Ceravolo, Scarabottolo, and Damiani (2016), Trcka and Pechenizkiy (2009) |
| Process mining to discover and check conformance of the students behavior within learning process (including self-regulated learning) Student Profile | Juhaňák, Zounek, and Rohlíková (2019), Intayoad, Kamyod, and Temdee (2018), Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, Morales, and Munoz-Gama (2018), Qiao and Hu (2018), Rodriguez, Nistal, Fonte, Penin, and Molina (2018), Cameranesi, Diamantini, Genga, and Potena (2017), Maldonado et al. (2017), Mukala, Cerone, and Turini (2017), Neyem, Diaz-Mosquera, Munoz-Gama, and Navon (2017) Alvarez, Fabra, Hernández, and Ezpeleta (2016), Aisa, Kurniati, and Yanuar Firdaus (2015), Beheshitha, Gašević, and Hatala (2015), Bannert, Reimann, and Sonnenberg (2013) Fernández-Gallego, Lama, Vidal, and Mucientes (2013), Martinovic, Drážd̆ilová, Slaninová, Kocyan, and Snasel (2012), Costa and Aparicio (2011), Malmberg, Järvelä, Järvenoja, and Panadero (2015), Premchaiswadi and Porouhan (2015a), Mittal and Sureka (2014) |
| Process mining to discover modeling patterns that can be associated with worse/better learning performance, improving guidance for conceptual modeling courses or reusing the model | Sedrakyan, Weerdt, and Snoeck (2016), Vidal, Vázquez-Barreiros, Lama, and Mucientes (2016), Sedrakyan, Snoeck, and Weerdt (2014) |
| Process mining to discover student's learning patterns for a specific educational activity (individual or collaborative) | Maita, Fantinato, Peres, Thom, and Hung (2017), Sobocinski, Malmberg, and Järvelä (2017), Wang, White, and Andersen (2017), Castillo (2016), Doleck, Jarrell, Poitras, Chaouachi, and Lajoie (2016), Monroy, Rangel, Bell, and Whitaker (2015), Premchaiswadi and Porouhan (2015b), Porouhan, Jongsawat, and Premchaiswadi (2014), Southavilay, Yacef, and Calvo (2010), Reimann, Frerejean, and Thompson (2009) |
| Process mining to discover students' behavior within learning process for assessment (including self-regulated learning) | Barreiros et al. (2014), Groba et al. (2014) |
| Recommendation of learning units based on discovered student's profile | Chen et al. (2015) |
| Repairing learning paths in adaptative e-learning systems based on process discovery | Blaskovic, Skopljanac-Macina, and Zakarija (2018) |
| Social network analysis to identify interaction patterns within a learning process | Porouhan and Premchaiswadi (2015), Premchaiswadi and Porouhan (2015b) |

### 5.5. Finance

This application domain focuses on the financial segment, e.g., banks, insurance companies, etc. Despite all process mining opportunities in auditing and the significant importance of this segment, we identified only 37 research papers. The papers cover different applications such as risk analysis, insurance claim handling, analysis of Automated Teller Machines (ATM) processes, validation of contractual security rules, gradual drift detection in profiles, financial auditing, fraud, and root cause analysis, improving performance of banking contact centers, loan approval, and credit card or clerk checks, among others.

Lakshmanan et al. (2013) proposed a probabilistic process model (PPM) using Markov methods that was applied in an automobile insurance company aimed to simulate insurance claims handling as a semi-structured business process. The proposed approach using combined process mining resulted in more accurate predictions than the traditional conditional probability analysis. In a large insurance company, Conforti, de Leoni, Rosa, van der Aalst, and ter Hofstede (2015) conducted research to equip a recommendation system focused on minimizing the overall risk in each instance considering faults and all usual process instances. This approach enables users to accept risk, while they are aware about decisions and impact when considering multiple process instances running concurrently. This application consist in a recommendation system is composed by: first, risk prediction using fault likelihoods and severity levels, in addition to individual statistics for each fault by considering the weight of each fault, which is calculated influence on the process; and; second, identifying the optimal performers to the task using integer linear programming

**Table 7**
Finance studies.

| Main Contribution | Studies |
|---|---|
| Data dependencies related to the accounting structure of recorded events used for discovering the process control-flow | Werner (2017) |
| Discovery of the data-flow process perspective for financial process | de Leoni and van der Aalst (2013) |
| Error detection and failure prediction in business processes based on event logs | Borkowski, Fdhila, Nardelli, Rinderle-Ma, and Schulte (2019) |
| Interval type-2 fuzzy sets in AHP (analytical hierarchy process) for weighting fraud cases detected using process mining (conformance checking) | Pane, Wibawa, and Purnomo (2016) |
| LTL checking and prediction based on decision-tree learning for checking goal achievement, false detection and oversight detection | Horita, Hirayama, Tahara, and Ohsuga (2016) |
| Method for process mining analysis on financial context | Huda, Ahmad, Sarno, and Santoso (2014), Kudo, Nogayama, Ishida, and Abe (2013), Weerdt, Schupp, Vanderloock, and Baesens (2013), Munoz-Gama and Echizen (2012) |
| Method for sudden and gradual drift detection for financial process | Maaradji et al. (2017) |
| Ontology-based process discovery to capture the business process anomalies using association rules | Sarno and Sinaga (2015) |
| Predicting future tasks using instance-specific probabilistic process model (PPM) | Lakshmanan et al. (2013) |
| Process discovery and conformance checking applied to fraud detection (auditing process) and/or fraud mitigation | Rahmawati, Sarno, Fatichah, and Sunaryono (2017), Rahmawati and Sarno (2017), Werner (2016), Werner and Gehrke (2015), Accorsi and Stocker (2012), Werner, Gehrke, and Nuttgens (2012), Accorsi, Wonnemann, and Dochow (2011), Jans, van der Werf, Lybaert, and Vanhoof (2011) |
| Process discovery and/or conformance checking applied on financial context | van Eck et al. (2017), Aldahami, Li, and Chan (2015), Arpasat, Porouhan, and Premchaiswadi (2015), de Leoni et al. (2014), Mahmood and Shaikh (2013), Peters, Dedene, and Poelmans (2013), Stuht, Speck, Feja, Witt, and Pulvermüller (2012), Wang, Yao, and Sun (2012), Awad, Weidlich, and Weske (2011), van Dongen, Jansen-Vullers, Verbeek, and van der Aalst (2007), Dustdar, Hoffmann, and van der Aalst (2005) |
| Process discovery, conformance checking and social network analysis applied to risk management | Caron, Vanthienen, and Baesens (2013a) |
| Process risk analysis by replaying event logs on a desired process model | Pika, van der Aalst, Wynn, Fidge, and ter Hofstede (2016) |
| Propagation of changes between aligned process models | Weidlich, Mendling, and Weske (2012) |
| Real-time monitoring of risks in executable business process models | Conforti et al. (2013) |
| Recommendation system for supporting risk-informed decisions based on process discovery (decision trees) and ILP (in case multiple process instances running concurrently) | Conforti et al. (2015) |
| Simulation system for operational decision support | Rozinat, Wynn, van der Aalst, ter Hofstede, and Fidge (2009) |

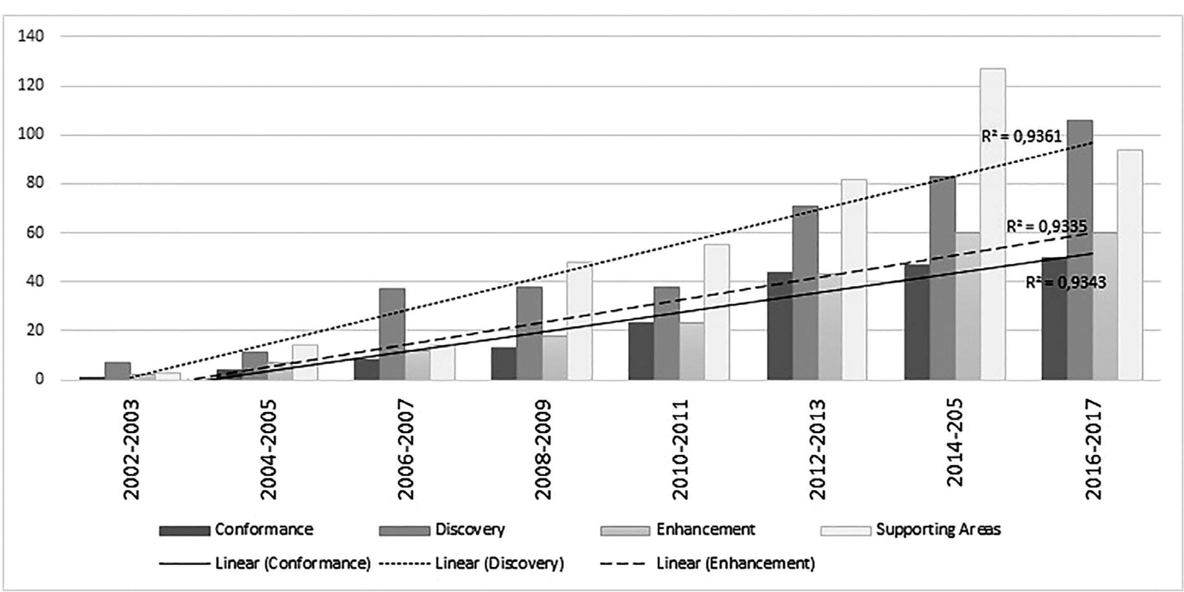


**Fig. 12.** Publications by year based on process mining types.

**Table 8**
Logistics studies.

| Main Contribution | Studies |
|---|---|
| Application of context awareness based on machine learning for improving process discovery | Becker and Intoyoad (2017) |
| Identification of design and operation problems for logistics process based on process mining techniques | Vasyutynskyy, Gellrich, Kabitzsch, and Wustmann (2010) |
| Multi-dimensional process mining approach | Sutrisnowati, Yahya, Bae, Pulshashi, and Adi (2017) |
| Prediction of event times based on a timed event graph with dynamic arc weights | Kecman and Goverde (2013) |
| Process discovery and social network analysis applied in a logistical context | Besri and Boulmakoul (2017), Li (2010) |
| Process discovery applied in a logistical context | Janssenswillen, Depaire, and Verboven (2017), Gou et al. (2016), Song, Jacobsen, Ye, and Ma (2016) Pulshashi, Bae, Sutrisnowati, Yahya, and Park (2015) Sutrisnowati et al. (2015), Sutrisnowati, Bae, Park, and Ha (2013), Krathu, Pichler, Zapletal, and Werthner (2012), Rozsnyai, Lakshmanan, Muthusamy, Khalaf, and Duftler (2012), Gellrich, Wagner, Vasyutynskyy, and Kabitzsch (2011), Haigh and Yaman (2011), Gerke et al. (2009), Gerke (2008), Gonzalez, Han, and Li (2006) |
| Process discovery, performance and conformance analyses for logistics processes | Wang, Caron, Vanthienen, Huang, and Guo (2014) |
| Process model discovery from event streams in a logistical context | Soffer et al. (2019), Repta and Stanescu (2017) |
| Remodeling logistic business processes based on process discovery | Wang, Zhu, Wang, and Huang (2016), Li and Deng (2009), Li (2009) |
| Resource allocation | Zhao, Zeng, Zheng, and Yang (2017) |
| Tools and techniques for analysing a logistics spaghetti process | Kumar, Thomas, and Annappa (2017) |

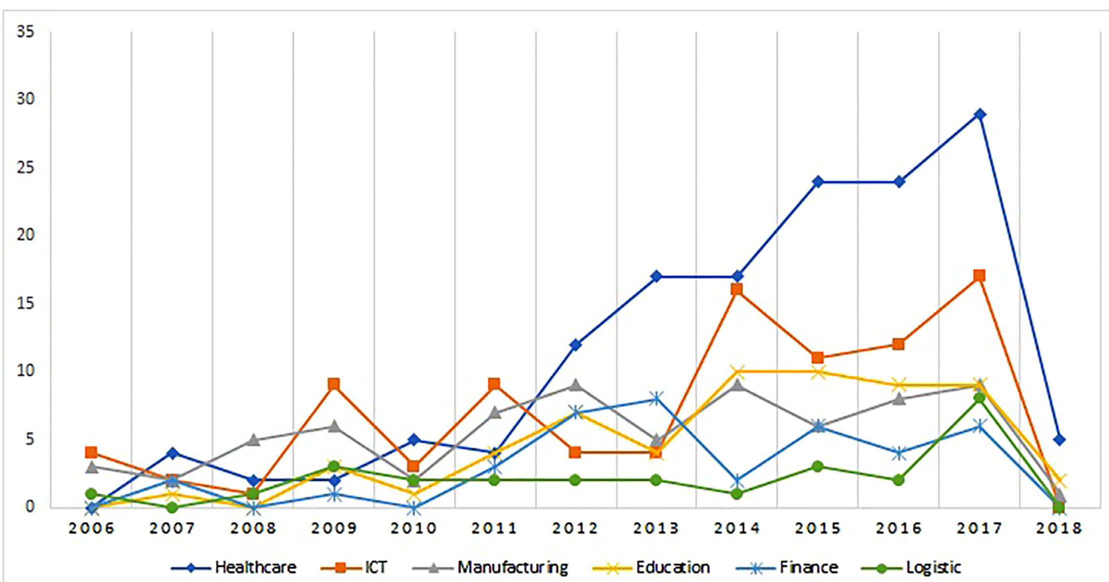


**Fig. 13.** Publications by year from top 6 application domains.

taking into account the predicted risk to distribute and assign activities. The result is also capable to optimize the global execution time.

An applied case was conducted by de Leoni, Suriadi, ter Hofstede, and van der Aalst (2014) that proposed a visualization tool for a large insurance company in Australia, called Suncorp$^{TM}$. The proposal focuses on a visualization framework able to produce a set of movies summarizing Suncorps claims. This movies contains performance and trends to provide the overall information to higher-level managers. The evaluation collected and analyzed over one million events with 34 activities, and it produced interesting maps containing a set of discoveries about types of losses, claimed payout, team allocation, and others.

In complex and unstructured process models, the analyses are difficult and time-consuming. Furthermore, there are often multiple views on the same process, so for this application an artifact-centric technique called CMS miner was proposed. The main goal was in obtaining correlations between artifact behavior to obtain insights and to check artifact interactions. It was applied to a loan process in a financial institution easily showing hidden information, and such insights easily point to issues in average time spent, and the top five most interesting artifact interactions, among others (van Eck, Sidorova, & van der Aalst, 2017).

Table 7 provide an overview of the process mining studies applied to the financial domain.

### 5.6. Logistics

In this study, we selected 27 studies focused on logistical processes regarding industrial shifts, warehouse layout improvements, bulk ports prediction, traffic control, detecting train rerouting, anomalies in freight, scheduling transport, and resource allocation in an aircraft company, among others.

Gerke, Claus, and Mendling (2009) used process mining to analyze events records in the supply chain system by SAP™. The challenge here was dealing with lack of data aggregation and difficulties in identifying process instances, because in logistic systems much information is not recorded as events. The work focused on different product shifts, such as: shipment aggregation, shipment disaggregation, product aggregation, transformation, and complex processing. The authors proposed the use of a radio frequency identification (RFID) system to identify the movement of goods in the logistics chain and process mining uses this information for analysis.

Sutrisnowati, Bae, and Song (2015) used process mining as a stage for the construction of Bayesian networks with the objective of analyzing the probability of delays in a port management system by considering container handling activities. The idea is to provide information to support port managers' decision-making regarding logistics activities.

Table 8 provide an overview of the process mining studies applied to the logistical domain.

### 5.7. Multiple application domains

Over 100 studies were published in domains such as: public municipal services for citizens, governmental layers; security services; call centers for customer services; usability and improvements for user experience; entertainment, sports and medias; high technological proposals focused on robotics, industry 4.0, and smart buildings; garments; advisory and audit offices; retail; biology sequences; hotels; and agriculture. The process mining applications are diverse and extendable to a wide range of businesses. This work is focused on peer reviewed papers; however, the software industry is starting to produce commercial materials to reach industries, thus producing business cases, product datasheets, and analyst opinion publications from research and advisory companies (e.g. Gartner[19]), among others.

## 6. Discussion

This systematic mapping study established research questions formulated in Section 2. With respect to RQ1: Which research topics can be identified in the primary studies of process mining? we started mapping the main types of process mining: discovery - focused on producing a real process model based on event logs; conformance - comparing a process model (*a priori* or discovered) and an event log, and vice versa; enhancement - to extend results with emphasis on frequencies, working time, waiting time, and bottlenecks. In addition to the main types, we suggested the supporting areas category, which group papers not directly related to one specific type of process mining but covering some crucial topics for implementing process mining projects or even showing already implemented process mining applications. Currently, process discovery is the most active type of process mining, also the supporting areas has been gaining greater attention and importance, as presented in Fig. 12. Also, the growth in the number of publications shows a strong uphill linear relationship for the three types of process mining. This confirms the growth of the process mining research area and its relevance for future research.

To make possible the application of the main types of process mining, we need some supporting areas. Six categories were defined toward this goal: process mining applications, methods for process mining projects, architecture and tools, background, gathering and cleaning and ontology (to extend event data). The three main types of process mining were mapped by splitting each one into more specific categories, mainly based on the concepts described by van der Aalst (2016) and the reading of the secondary papers. Regarding question (RQ2), we identified hundreds of algorithms, but it was possible to map the most popular, for which improvements are usually suggested, as well as a set of applications using these algorithms. As the most used algorithms we identified Heuristic Miner, Fuzzy Miner, Alpha miner, and more recently Inductive Miner. Finally, regarding RQ3, we evaluated where process mining has been applied. We identified that process mining has already been used among different application domains and its applications continue to increase, as presented in Fig. 13.

## 7. Conclusion and future work

The research method applied in this work delivered a breadth-first review of the primary research studies in process mining and provided an overview of the published applications domain. Evaluation of 3713 published papers over the last 16 years resulted in the selection of more than 34% of the papers related to process mining. This large base allowed us to produce a comprehensive map regarding the established research questions.

When studying a research area, it is relevant to identify the research topics, e.g., subjects receiving more attention, as well as beginning or trending topics, the accepting journals and conferences, the most active universities, authors, and so forth. The present paper analyzed various process mining research topics starting with the process discovery, conformance, and enhancement. Further, supporting areas were identified as process mining applications, methods for process mining projects, architecture and tools, background, gathering and cleaning, and ontology. The second research question was related to mapping the algorithms used for process discovery. The question explored hundreds of proposals, where we noted promising algorithms or algorithm families, such as: heuristic miner, fuzzy miner, Alpha miner, Markov-based, based on evolutionary algorithms, inductive families, clustering-based, and ILP-based, among others. For the third and last question, application domains were reviewed. The segments of Healthcare, ICT, Manufacturing, Education, Finance, and Logistics are the most active. However, process mining applications are not limited to these domains and it is possible to extend these to almost any business process.

As further work, there are several possibilities such as process variabilities control using similarities measurements involving: state space, natural language, graph structures, behavior-based, performance level, cost groups, or even based on wizards providing expert opinions. There are also many open gaps for infrastructure improvements related to performance and scalability, such as combining pipelines for event data processing, and evolution of algorithms for parallelization. The big data technology companies and projects are not focused on process mining (van der Aalst & Damiani, 2015); consequently, few research efforts related process mining applied on big data scenarios and applications with large event logs. Applying and scaling process mining using big data pipelines and stream processing are a promising research subject. Many algorithms implemented in the ProM framework have high computational cost, and real applications require more effort in their optimization.

Process discovery algorithms usually work based on one-class induction. Very few approaches are capable of discovery considering multi-classes based on a context, outcome, or process variant. A second common limitation of the discovery algorithms is related to the Petri net language bias, caused by structural limitations and complexity for transforming exclusive OR gates to inclusive OR. Few discovery algorithms are based on other languages.

[19] https://www.gartner.com/.

Another high-visibility sub-area but with sparse research efforts is the replay. An approach starting from replay that establishes a model to begin simulation and projection should prove interesting for further research. Other opportunities are related to offering process mining advancements for producing more intelligent approaches by projecting performance, predicating decisions, predicting process outcomes, trending KPI's impact, combining alarms for anomaly detection, drift detection, and process improvements recommendations.

The business processes control different areas and aspects of the companies, especially when they need to be collaborative, which is essential on today business. Process mining can provide a non-empirical method for process analyzes and optimizations anchored on how things are working for real. This is essential to provide more control for collaborative business processes when it is necessary to aggregate information about processes involving different resources, departments, etc. The classical business process management (BPM) approaches usually takes a top-down approach, starting by designing the processes in a high-level model. However, in collaborative scenarios, this is expensive and susceptible to errors. A bottom-up approach, such as process mining, can discover and also enhance these models based on evidence and with a certain confidence degree, contributing to the different business process analyses.

Although there are a variety of practical analysis techniques that can be applied to a discovered business model, they present some limitation because the event log maybe not directly linked to any concepts. On this situation, the event log can be enhanced by including some semantics, based on an ontology, in order to improve the business process analyzes, as described by Kingsley et al. (2016) and de Medeiros and van der Aalst (2008). An ontological model can better structure performance knowledge, also enrich the evaluation of the business process models considering different aspects, improving assessments or even tendency analyzes.

One of the most promising areas for the use of process mining is the industrial area, due to Industry 4.0. Industry 4.0 is considered as the process of transforming traditional production systems into Cyber-Physical Systems (CPS) (as cited in Liao, Deschamps, de Freitas Rocha Loures, & Ramos, 2017, and Xu, Xu, & Li, 2018). Connectivity and the immense amount of information that characterizes Industry 4.0 together with data analysis tools are the basis for the smart factory concept (as cited in Patel, Ali, & Sheth, 2018). Several techniques and data analysis tools have emerged as one of the main pillars of Industry 4.0 (as cited in Liao et al., 2017, and Xu et al., 2018). All the digitization of the factory floor with Industry 4.0 has enabled a new level of self-organization and optimization of manufacturing operations. In this context, smart products and smart services have emerged as a result of the large amount of data and analytical techniques available. For example, manufacturers or service providers can make maintenance-related forecasts based on real-time data on equipment wear (predictive maintenance). And the equipment itself can self-manage in relation to their respective maintenance and shutdowns, through requisitions of items to be replaced, allocation of teams, among others. Thus, Industry 4.0 is emerging as an important driver of process mining.

## Conflict of Interest

All authorscertify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## Acknowledgement

The authors would like to thank for supporting by CAPES *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* under Grant No.:88887.147294/2017-00, 88887.147202/2017-00, and 88887.321450/2019-00. This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - Brasil (CAPES) - Finance Code 001.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.eswa.2019.05.003

## Credit authorship contribution statement

**Cleiton dos Santos Garcia:** Methodology, Writing - original draft, Project administration. **Alex Meincheim:** Formal analysis. **Elio Ribeiro Faria Junior:** Formal analysis, Writing - review & editing. **Marcelo Rosano Dallagassa:** Formal analysis. **Denise Maria Vecino Sato:** Formal analysis. **Deborah Ribeiro Carvalho:** Formal analysis. **Eduardo Alves Portela Santos:** Methodology, Formal analysis. **Edson Emilio Scalabrin:** Project administration, Formal analysis, Writing - original draft.

## References


Anonymous (2016). IEEE standard for eXtensible Event Stream (XES) for achieving interoperability in event logs and event streams. *IEEE Std 1849–2016*, 1–50. doi:10.1109/IEEESTD.2016.7740858.

van der Aalst, W. (2011). Process mining: discovering and improving spaghetti and lasagna processes. *2011 IEEE symposium on computational intelligence and data mining (CIDM)*. IEEE. doi:10.1109/cidm.2011.6129461.

van der Aalst, W. (2012a). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems, 3*(2), 1–17. doi:10.1145/2229156.2229157.

van der Aalst, W. M. P. (2012b). What makes a good process model? *Software & Systems Modeling, 11*(4), 557–569. doi:10.1007/s10270-012-0265-9.

van der Aalst, W. (2016). *Process mining: Data science in action*. Springer Berlin Heidelberg. doi:10.1007/978-3-662-49851-4.

van der Aalst, W., Adriansyah, A., de Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., et al. (2012). Process mining manifesto. In *Business process management workshops on process mining* (pp. 169–194). Springer Berlin Heidelberg. doi:10.1007/978-3-642-28108-2_19.

van der Aalst, W., & Damiani, E. (2015). Processes meet big data: Connecting data science with process science. *IEEE Transactions on Services Computing, 8*(6), 810–819. doi:10.1109/tsc.2015.2493732.

van der Aalst, W., van Dongen, B., Herbst, J., Maruster, L., Schimm, G., & Weijters, A. (2003). Workflow mining: A survey of issues and approaches. *Data & Knowledge Engineering, 47*(2), 237–267. doi:10.1016/s0169-023x(03)00066-1.

van der Aalst, W., Reijers, H., Weijters, A., van Dongen, B., de Medeiros, A. A., Song, M., et al. (2007). Business process mining: An industrial application. *Information Systems, 32*(5), 713–732. doi:10.1016/j.is.2006.05.003.

van der Aalst, W., & Weijters, A. (2004). Process mining: A research agenda. *Computers in Industry, 53*(3), 231–244. doi:10.1016/j.compind.2003.10.001.

van der Aalst, W., Weijters, A., & Maruster, L. (2002). Workflow mining: Which processes can be rediscovered? In *Eindhoven university of technology* (pp. 1–25).

van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering, 16*(9), 1128–1142. doi:10.1109/TKDE.2004.47.

van der Aalst, W. M. (2006). Matching observed behavior and modeled behavior: An approach based on petri nets and integer programming. *Decision Support Systems, 42*(3), 1843–1859. doi:10.1016/j.dss.2006.03.013.

van der Aalst, W. M. P., de Medeiros, A. K. A., & Weijters, A. J. M. M. (2005). Genetic process mining. In *Applications and theory of petri nets 2005* (pp. 48–69). Springer Berlin Heidelberg. doi:10.1007/11494744_5.

van der Aalst, W. M. P., Reijers, H. A., & Song, M. (2005). Discovering social networks from event logs. *Computer Supported Cooperative Work (CSCW), 14*(6), 549–593. doi:10.1007/s10606-005-9005-9.

Aalst, W. V. D. (2001). Daelemans. automated discovery of workflow models from hospital data. In *Proceedings 13th Belgium-Netherlands conference on artificial intelligence (BNAIC'01* (pp. 25–26).

Aalst, W. v. d. (2015). Big software on the run: In vivo software analytics based on process mining (keynote). In *Proceedings of the 2015 international conference on software and system process*. In *ICSSP 2015* (pp. 1–5). New York, NY, USA: ACM. doi:10.1145/2785592.2785593.

Abohamad, W., Ramy, A., & Arisha, A. (2017). A hybrid process-mining approach for simulation modeling. *2017 winter simulation conference (WSC)*. IEEE. doi:10.1109/wsc.2017.8247894.

Acampora, G., Bernardi, M. L., Cimitile, M., Tortora, G., & Vitiello, A. (2017). A fuzzy-based autoscaling approach for process centered cloud systems. *2017 IEEE international conference on fuzzy systems (FUZZ-IEEE)*. IEEE. doi:10.1109/fuzz-ieee.2017.8015768.

Accorsi, R., & Stocker, T. (2012). On the exploitation of process mining for security audits. In *Proceedings of the 27th annual ACM symposium on applied computing - SAC '12*. ACM Press. doi:10.1145/2245276.2232051.

Accorsi, R., Wonnemann, C., & Dochow, S. (2011). SWAT: A security workflow analysis toolkit for reliably secure process-aware information systems. *2011 sixth international conference on availability, reliability and security*. IEEE. doi:10.1109/ares.2011.108.

Agrawal, R., Gunopulos, D., & Leymann, F. (1998). Mining process models from workflow logs. In H.-J. Schek, G. Alonso, F. Saltor, & I. Ramos (Eds.), *Advances in database technology — EDBT'98* (pp. 467–483). Berlin, Heidelberg: Springer Berlin Heidelberg.

Aguirre, S., Parra, C., & Alvarado, J. (2013). Combination of process mining and simulation techniques for business process redesign: A methodological approach. In *Lecture notes in business information processing* (pp. 24–43). Springer Berlin Heidelberg. doi:10.1007/978-3-642-40919-6_2.

Aisa, V., Kurniati, A. P., & Yanuar Firdaus, A. W. (2015). Evaluation of the online assessment test using process mining (case study: Intensive english center). In *2015 3rd international conference on information and communication technology (ICOICT)* (pp. 472–477). doi:10.1109/ICoICT.2015.7231470.

Akman, B., & Demirörs, O. (2009). Applicability of process discovery algorithms for software organizations. *2009 35th euromicro conference on software engineering and advanced applications*. IEEE. doi:10.1109/seaa.2009.87.

Aksu, U., Schunselaar, D., & Reijers, H. (2016). A cross-organizational process mining framework for obtaining insights from software products: Accurate comparison challenges. *2016 IEEE 18th conference on business informatics (CBI)*. IEEE. doi:10.1109/cbi.2016.25.

Aldahami, A., Li, Y., & Chan, T. (2015). Using a flow graph to represent data flow and dependency in event logs. *2015 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology (WI-IAT)*. IEEE. doi:10.1109/wi-iat.2015.213.

Alizadeh, M., Lu, X., Fahland, D., Zannone, N., & van der Aalst, W. M. (2018). Linking data and process perspectives for conformance analysis. *Computers & Security, 73*, 172–193. doi:10.1016/j.cose.2017.10.010.

Alonso, I. G., Fuente, M. P. A. G., & Brugos, J. A. L. (2009). Using sysml to describe a new methodology for semiautomatic software generation from inferred behavioral and data models. *2009 fourth international conference on systems*. IEEE. doi:10.1109/icons.2009.50.

Alvarez, C., Rojas, E., Arias, M., Munoz-Gama, J., Sepúlveda, M., Herskovic, V., et al. (2018). Discovering role interaction models in the emergency room using process mining. *Journal of Biomedical Informatics, 78*, 60–77. doi:10.1016/j.jbi.2017.12.015.

Alvarez, P., Fabra, J., Hernández, S., & Ezpeleta, J. (2016). Alignment of teacher's plan and students' use of lms resources. analysis of moodle logs. In *2016 15th international conference on information technology based higher education and training (ithet)* (pp. 1–8). doi:10.1109/ITHET.2016.7760720.

Anuwatvisit, S., Tungkasthan, A., & Premchaiswadi, W. (2012). Bottleneck mining and petri net simulation in education situations. In *2012 tenth international conference on ict and knowledge engineering* (pp. 244–251). doi:10.1109/ICTKE.2012.6408562.

Ariouat, H., Cairns, A. H., Barkaoui, K., Akoka, J., & Khelifa, N. (2016). A two-step clustering approach for improving educational process model discovery. In *2016 ieee 25th international conference on enabling technologies: Infrastructure for collaborative enterprises (wetice)* (pp. 38–43). doi:10.1109/WETICE.2016.18.

Arpasat, P., Porouhan, P., & Premchaiswadi, W. (2015). Improvement of call center customer service in a thai bank using disco fuzzy mining algorithm. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368477.

Association of Certified Fraud Examiners (2016). Report to the nation on occupational fraud & abuse. *Global Fraud Study*, 1–92.

Astromskis, S., Janes, A., & Mahdiraji, A. R. (2012). Egidio: A non-invasive approach for synthesizing organizational models. *2012 34th international conference on software engineering (ICSE)*. IEEE. doi:10.1109/icse.2012.6227062.

Atastina, I., & Kurniati, A. P. (2014). Student registration process evaluation using process mining case study: IT telkom. *Ninth international conference on digital information management (ICDIM 2014)*. IEEE. doi:10.1109/icdim.2014.6991417.

Augusto, V., Xie, X., Prodel, M., Jouaneton, B., & Lamarsalle, L. (2016). Evaluation of discovered clinical pathways using process mining and joint agent-based discrete-event simulation. *2016 winter simulation conference (WSC)*. IEEE. doi:10.1109/wsc.2016.7822256.

Awad, A., Weidlich, M., & Weske, M. (2011). Visually specifying compliance rules and explaining their violations for business processes. *Journal of Visual Languages & Computing, 22*(1), 30–55. doi:10.1016/j.jvlc.2010.11.002.

Awad, A., Zaki, N. M., & Francescomarino, C. D. (2016). Analyzing and repairing overlapping work items in process logs. *Information and Software Technology, 80*, 110–123. doi:10.1016/j.infsof.2016.08.010.

Ayutaya, N. S. N., Palungsuntikul, P., & Premchaiswadi, W. (2012). Heuristic mining: Adaptive process simplification in education. In *2012 tenth international conference on ict and knowledge engineering* (pp. 221–227). doi:10.1109/ICTKE.2012.6408559.

Azzini, A., Ceravolo, P., Scarabottolo, N., & Damiani, E. (2016). On the predictive power of university curricula. In *2016 IEEE global engineering education conference (educon)* (pp. 929–932). doi:10.1109/EDUCON.2016.7474663.

Baier, T., Mendling, J., & Weske, M. (2014). Bridging abstraction layers in process mining. *Information Systems, 46*, 123–139. doi:10.1016/j.is.2014.04.004.

Baker, K., Dunwoodie, E., Jones, R. G., Newsham, A., Johnson, O., Price, C. P., et al. (2017). Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. *International Journal of Medical Informatics, 103*, 32–41. doi:10.1016/j.ijmedinf.2017.03.011.

Bannert, M., Reimann, P., & Sonnenberg, C. (2013). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. doi:10.1007/s11409-013-9107-6.

Barreiros, B. V., Lama, M., Mucientes, M., & Vidal, J. C. (2014). Softlearn: A process mining platform for the discovery of learning paths. In *2014 IEEE 14th international conference on advanced learning technologies* (pp. 373–375). doi:10.1109/ICALT.2014.111.

Bartík, V., & Pospíšil, M. (2015). Use of frequent itemset mining techniques to analyze business processes. In *2015 7th international joint conference on knowledge discovery, knowledge engineering and knowledge management (ic3k): 01* (pp. 273–280).

Basole, R. C., Braunstein, M. L., Kumar, V., Park, H., Kahng, M., Chau, D. H. P., et al. (2015). Understanding variations in pediatric asthma care processes in the emergency department using visual analytics. *Journal of the American Medical Informatics Association, 22*(2), 318–323. doi:10.1093/jamia/ocu016.

Basole, R. C., Park, H., Gupta, M., Braunstein, M. L., Chau, D. H., & Thompson, M. (2015). A visual analytics approach to understanding care process variation and conformance. In *Proceedings of the 2015 workshop on visual analytics in healthcare - VAHC '15*. ACM Press. doi:10.1145/2836034.2836040.

Becker, T., & Intoyoad, W. (2017). Context aware process mining in logistics. *Procedia CIRP, 63*, 557–562. doi:10.1016/j.procir.2017.03.149.

Beheshitha, S. S., Gašević, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. In *Proceedings of the fifth international conference on learning analytics and knowledge - LAK '15*. ACM Press. doi:10.1145/2723576.2723628.

Benner-Wickner, M., Brückmann, T., Gruhn, V., & Book, M. (2015). Process mining for knowledge-intensive business processes. In *Proceedings of the 15th international conference on knowledge technologies and data-driven business - i-KNOW '15*. ACM Press. doi:10.1145/2809563.2809580.

Bergenthum, R., Desel, J., Lorenz, R., & Mauser, S. (2007a). Process mining based on regions of languages. In *Lecture notes in computer science* (pp. 375–383). Springer Berlin Heidelberg. doi:10.1007/978-3-540-75183-0_27.

Bergenthum, R., Desel, J., Lorenz, R., & Mauser, S. (2007b). Process mining based on regions of languages. In *Lecture notes in computer science* (pp. 375–383). Springer Berlin Heidelberg. doi:10.1007/978-3-540-75183-0_27.

Berlingerio, M., Pinelli, F., Nanni, M., & Giannotti, F. (2009). Temporal mining for interactive workflow data analysis. In *Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining - KDD '09*. ACM Press. doi:10.1145/1557019.1557038.

Bernardi, M. L., Cimitile, M., Francescomarino, C. D., & Maggi, F. M. (2016). Do activity lifecycles affect the validity of a business rule in a business process? *Information Systems, 62*, 42–59. doi:10.1016/j.is.2016.06.002.

Bernardi, S., Alastuey, R. P., & Trillo-Lado, R. (2017). Using process mining and model-driven engineering to enhance security of web information systems. *2017 IEEE european symposium on security and privacy workshops (EuroS&PW)*. IEEE. doi:10.1109/eurospw.2017.66.

Berriche, F.-Z., Zeddini, B., Kadima, H., & Riviere, A. (2015). Combining case-based reasoning and process mining to improve collaborative decision-making in products design. *2015 IEEE/ACS 12th international conference of computer systems and applications (AICCSA)*. IEEE. doi:10.1109/aiccsa.2015.7507199.

Besri, Z., & Boulmakoul, A. (2017). Framework for organizational structure re-design by assessing logistics' business processes in harbor container terminals. *Transportation Research Procedia, 22*, 164–173. doi:10.1016/j.trpro.2017.03.023.

Biermann, A. W., & Feldman, J. A. (1972). On the synthesis of finite-state machines from samples of their behavior. *IEEE Transactions on Computers, C-21*(6), 592–597. doi:10.1109/tc.1972.5009015.

Blaskovic, B., Skopljanac-Macina, F., & Zakarija, I. (2018). Discovering e-learning process models from counterexamples. *2018 41st international convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE. doi:10.23919/mipro.2018.8400112.

Blum, F. R. (2016). Mining software process lines. In *Proceedings of the 38th international conference on software engineering companion - ICSE '16*. ACM Press. doi:10.1145/2889160.2889267.

Blum, F. R., Simmonds, J., & Bastarrica, M. C. (2015). Software process line discovery. In *Proceedings of the 2015 international conference on software and system process - ICSSP 2015*. ACM Press. doi:10.1145/2785592.2785605.

Blum, T., Padoy, N., Feußner, H., & Navab, N. (2008). Modeling and online recognition of surgical phases using hidden markov models. In *Medical image computing and computer-assisted intervention – MICCAI 2008* (pp. 627–635). Springer Berlin Heidelberg. doi:10.1007/978-3-540-85990-1_75.

Bobek, S., & Nalepa, G. J. (2017). Uncertain context data management in dynamic mobile environments. *Future Generation Computer Systems, 66*, 110–124. doi:10.1016/j.future.2016.06.007.

Bogarín, A., Romero, C., Cerezo, R., & Sánchez-Santillán, M. (2014). Clustering for improving educational process mining. *Proceedins of the fourth international conference on learning analytics and knowledge - LAK '14*. ACM Press. doi:10.1145/2567574.2567604.

Borkowski, M., Fdhila, W., Nardelli, M., Rinderle-Ma, S., & Schulte, S. (2019). Event-based failure prediction in distributed business processes. *Information Systems, 81*, 220–235. doi:10.1016/j.is.2017.12.005.

Bose, R. J. C., Mans, R. S., & van der Aalst, W. M. (2013). Wanna improve process mining results? In *2013 IEEE symposium on computational intelligence and data mining (CIDM)* (pp. 127–134). IEEE. doi:10.1109/CIDM.2013.6597227.

Bose, R. P. J. C., van der Aalst, W. M. P., Zliobaite, I., & Pechenizkiy, M. (2014). Dealing with concept drifts in process mining. *IEEE Transactions on Neural Networks and Learning Systems, 25*(1), 154–171. doi:10.1109/TNNLS.2013.2278313.

Bouarfa, L., & Dankelman, J. (2012). Workflow mining and outlier detection from clinical activity logs. *Journal of Biomedical Informatics, 45*(6), 1185–1190. doi:10.1016/j.jbi.2012.08.003.

vanden Broucke, S. K., & Weerdt, J. D. (2017). Fodina: A robust and flexible heuristic process discovery technique. *Decision Support Systems, 100*, 109–118. doi:10.1016/j.dss.2017.04.005.

Bruntsch, A., & Tseng, M. M. (2016). An approach for process control of responsive service processes. *CIRP Annals, 65*(1), 459–462. doi:10.1016/j.cirp.2016.04.022.

Burattin, A., Sperduti, A., & Veluscek, M. (2013). Business models enhancement through discovery of roles. *2013 IEEE symposium on computational intelligence and data mining (CIDM)*. IEEE. doi:10.1109/cidm.2013.6597224.

Cadez, I., Heckerman, D., Meek, C., Smyth, P., & White, S. (2000). Visualization of navigation patterns on a web site using model-based clustering. In *Proceedings of the sixth ACM SIGKDD international conference on knowledge discovery and data mining - KDD '00*. ACM Press. doi:10.1145/347090.347151.

Cairns, A. H., Gueni, B., Hafdi, H., Joubert, C., & Khelifa, N. (2015). Towards a distributed computation platform tailored for educational process discovery and analysis. *2015 international conference on protocol engineering (ICPE) and international conference on new technologies of distributed systems (NTDS)*. IEEE. doi:10.1109/notere.2015.7293494.

Calvo, R. A., O'Rourke, S. T., Jones, J., Yacef, K., & Reimann, P. (2011). Collaborative writing support tools on the cloud. *IEEE Transactions on Learning Technologies, 4*(1), 88–97. doi:10.1109/TLT.2010.43.

Cameranesi, M., Diamantini, C., Genga, L., & Potena, D. (2017). Students' careers analysis. In *Proceedings of the 7th international conference on web intelligence, mining and semantics - WIMS '17*. ACM Press. doi:10.1145/3102254.3102270.

Caron, F., Vanthienen, J., & Baesens, B. (2013a). A comprehensive investigation of the applicability of process mining techniques for enterprise risk management. *Computers in Industry, 64*(4), 464–475. doi:10.1016/j.compind.2013.02.001.

Caron, F., Vanthienen, J., & Baesens, B. (2013b). Healthcare analytics: Examining the diagnosis–treatment cycle. *Procedia Technology, 9*, 996–1004. doi:10.1016/j.protcy.2013.12.111.

Caron, F., Vanthienen, J., Vanhaecht, K., Limbergen, E. V., De Weerdt, J., & Baesens, B. (2014a). A process mining-based investigation of adverse events in care processes. *Health Information Management Journal, 43*(1), 16–25. doi:10.1177/183335831404300103.

Caron, F., Vanthienen, J., Vanhaecht, K., Limbergen, E. V., De Weerdt, J. D., & Baesens, B. (2014b). Monitoring care processes in the gynecologic oncology department. *Computers in Biology and Medicine, 44*, 88–96. doi:10.1016/j.compbiomed.2013.10.015.

Castillo, L. (2016). A virtual laboratory for multiagent systems: Joining efficacy, learning analytics and student satisfaction. In *2016 international symposium on computers in education (siie)* (pp. 1–6). doi:10.1109/SIIE.2016.7751820.

Cecconi, A., Ciccio, C. D., Giacomo, G. D., & Mendling, J. (2018). Interestingness of traces in declarative process mining: The janus LTLp$$_f$$ approach. In *Lecture notes in computer science* (pp. 121–138). Springer International Publishing. doi:10.1007/978-3-319-98648-7_8.

Ceravolo, P., Azzini, A., Damiani, E., Lazoi, M., Marra, M., & Corallo, A. (2016). Translating process mining results into intelligible business information. In *Proceedings of the 11th international knowledge management in organizations conference on the changing face of knowledge management impacting society - KMO '16*. ACM Press. doi:10.1145/2925995.2925997.

Chabrol, M., Dalmas, B., Norre, S., & Rodier, S. (2016). A process tree-based algorithm for the detection of implicit dependencies. *2016 IEEE tenth international conference on research challenges in information science (RCIS)*. IEEE. doi:10.1109/rcis.2016.7549357.

Chen, J., Zhang, Y., Sun, J., Chen, Y., Lin, F., & Jin, Q. (2015). Personalized micro-learning support based on process mining. *2015 7th international conference on information technology in medicine and education (ITME)*. IEEE. doi:10.1109/itme.2015.120.

Chen, N., Hoi, S. C. H., & Xiao, X. (2011). Software process evaluation: A machine learning approach. *2011 26th IEEE/ACM international conference on automated software engineering (ASE 2011)*. IEEE. doi:10.1109/ase.2011.6100070.

Chen, S., Yang, S., Zhou, M., Burd, R., & Marsic, I. (2017). Process-oriented iterative multiple alignment for medical process mining. *2017 IEEE international conference on data mining workshops (ICDMW)*. IEEE. doi:10.1109/icdmw.2017.63.

Chen, Y., Kho, A. N., Liebovitz, D., Ivory, C., Osmundson, S., Bian, J., et al. (2018). Learning bundled care opportunities from electronic medical records. *Journal of Biomedical Informatics, 77*, 1–10. doi:10.1016/j.jbi.2017.11.014.

Cheng, H. J., Ou-Yang, C., & Juan, Y. C. (2012). A hybrid evolutionary-based process mining technology to discover parallelism structures. *2012 IEEE international conference on industrial engineering and engineering management*. IEEE. doi:10.1109/ieem.2012.6838011.

Chinces, D., & Salomie, I. (2015). Optimizing spaghetti process models. *2015 20th international conference on control systems and computer science*. IEEE. doi:10.1109/cscs.2015.15.

Chindenga, E., Scott, M. S., & Gurajena, C. (2017). Semantics based service orchestration in IoT. In *Proceedings of the South African institute of computer scientists and information technologists on - SAICSIT '17*. ACM Press. doi:10.1145/3129416.3129438.

Cho, M., Song, M., Comuzzi, M., & Yoo, S. (2017). Evaluating the effect of best practices for business process redesign: An evidence-based approach based on process mining techniques. *Decision Support Systems, 104*, 92–103. doi:10.1016/j.dss.2017.10.004.

Cho, M., Song, M., & Yoo, S. (2014). A systematic methodology for outpatient process analysis based on process mining. In C. Ouyang, & J.-Y. Jung (Eds.), *Asia pacific business process management* (pp. 31–42). Cham: Springer International Publishing.

Chomyat, W., & Premchaiswadi, W. (2016). Process mining on medical treatment history using conformance checking. *2016 14th international conference on ICT and knowledge engineering (ICT&KE)*. IEEE. doi:10.1109/ictke.2016.7804102.

Ciccarese, P., Caffi, E., Boiocchi, L., Halevy, A., Quaglini, S., Kumar, A., et al. (2003). The NewGuide project: Guidelines, information sharing and learning from exceptions. In *Artificial intelligence in medicine* (pp. 163–167). Springer Berlin Heidelberg. doi:10.1007/978-3-540-39907-0_23.

Colombo, A., Damiani, E., & Gianini, G. (2006). Discovering the software process by means of stochastic workflow analysis. *Journal of Systems Architecture, 52*(11), 684–692. doi:10.1016/j.sysarc.2006.06.012.

Compagna, L., dos Santos, D. R., Ponta, S. E., & Ranise, S. (2017). Aegis: Automatic enforcement of security policies in workflow-driven web applications. In *Proceedings of the seventh ACM on conference on data and application security and privacy - CODASPY '17*. ACM Press. doi:10.1145/3029806.3029813.

Comuzzi, M., Vonk, J., & Grefen, P. (2012). Measures and mechanisms for process monitoring in evolving business networks. *Data & Knowledge Engineering, 71*(1), 1–28. doi:10.1016/j.datak.2011.07.004.

Conforti, R., de Leoni, M., Rosa, M. L., van der Aalst, W. M., & ter Hofstede, A. H. (2015). A recommendation system for predicting risks across multiple business process instances. *Decision Support Systems, 69*, 1–19. doi:10.1016/j.dss.2014.10.006.

Conforti, R., Rosa, M. L., Fortino, G., ter Hofstede, A. H., Recker, J., & Adams, M. (2013). Real-time risk monitoring in business processes: A sensor-based approach. *Journal of Systems and Software, 86*(11), 2939–2965. doi:10.1016/j.jss.2013.07.024.

Cook, J. E., & Wolf, A. L. (1995). Automating process discovery through event-data analysis. In *Proceedings of the 17th international conference on software engineering - ICSE '95*. ACM Press. doi:10.1145/225014.225021.

Cook, J. E., & Wolf, A. L. (1998). Discovering models of software processes from event-based data. *ACM Transactions on Software Engineering and Methodology, 7*(3), 215–249. doi:10.1145/287000.287001.

Costa, C. J., & Aparicio, M. (2011). Analysis of e-learning processes. In *Proceedings of the 2011 workshop on open source and design of communication - OSDOC '11*. ACM Press. doi:10.1145/2016716.2016726.

da Costa Cordeiro, W. L., Machado, G. S., Andreis, F. G., Wickboldt, J. A., Lunardi, R. C., dos Santos, A. D., et al. (2009). CHANGEMINER: A solution for discovering IT change templates from past execution traces. *2009 IFIP/IEEE international symposium on integrated network management*. IEEE. doi:10.1109/inm.2009.5188792.

Curia, R., Gallucci, L., & Ruffolo, M. (2005). Knowledge management in health care: An architectural framework for clinical process management systems. *16th international workshop on database and expert systems applications (DEXA'05)*. IEEE. doi:10.1109/dexa.2005.125.

Dabrowski, J., Kifetew, F. M., Munante, D., Letier, E., Siena, A., & Susi, A. (2017). Discovering requirements through goal-driven process mining. *2017 IEEE 25th international requirements engineering conference workshops (REW)*. IEEE. doi:10.1109/rew.2017.61.

Dagliati, A., Sacchi, L., Cerra, C., Leporati, P., Cata, P. D., Chiovato, L., et al. (2014). Temporal data mining and process mining techniques to identify cardiovascular risk-associated clinical pathways in type 2 diabetes patients. *IEEE-EMBS international conference on biomedical and health informatics (BHI)*. IEEE. doi:10.1109/bhi.2014.6864348.

Dagliati, A., Sacchi, L., Zambelli, A., Tibollo, V., Pavesi, L., Holmes, J., et al. (2017). Temporal electronic phenotyping by mining careflows of breast cancer patients. *Journal of Biomedical Informatics, 66*, 136–147. doi:10.1016/j.jbi.2016.12.012.

Damevski, K., Chen, H., Shepherd, D., & Pollock, L. (2016). Interactive exploration of developer interaction traces using a hidden markov model. In *Proceedings of the 13th international workshop on mining software repositories - MSR '16*. ACM Press. doi:10.1145/2901739.2901741.

Damevski, K., Shepherd, D. C., Schneider, J., & Pollock, L. (2017). Mining sequences of developer interactions in visual studio for usage smells. *IEEE Transactions on Software Engineering, 43*(4), 359–371. doi:10.1109/tse.2016.2592905.

Darabi, H., Galanter, W. L., Lin, J. Y.-Y., Buy, U., & Sampath, R. (2009). Modeling and integration of hospital information systems with petri nets. *2009 IEEE/INFORMS international conference on service operations, logistics and informatics*. IEEE. doi:10.1109/soli.2009.5203928.

Das, S., & Mozer, M. C. (1994). A unified gradient-descent/clustering architecture for finite state machine induction. In *Nips* (pp. 19–26). Morgan Kaufmann.

Dasani, S., Shah, S. L., Chen, T., Funnell, J., & Pollard, R. W. (2015). Monitoring safety of process operations using industrial workflows. *IFAC-PapersOnLine, 48*(8), 451–456. doi:10.1016/j.ifacol.2015.09.009.

De Smedt, J., De Weerdt, J., & Vanthienen, J. (2014). Multi-paradigm process mining: Retrieving better models by combining rules and sequences. In R. Meersman, H. Panetto, T. Dillon, M. Missikoff, L. Liu, O. Pastor, & T. Sellis (Eds.), *On the move to meaningful internet systems: Otm 2014 conferences* (pp. 446–453). Berlin, Heidelberg: Springer Berlin Heidelberg.

Defossez, G., Rollet, A., Dameron, O., & Ingrand, P. (2014). Temporal representation of care trajectories of cancer patients using data from a regional information system: an application in breast cancer. *BMC Medical Informatics and Decision Making, 14*(1). doi:10.1186/1472-6947-14-24.

Delgado, A., Weber, B., Ruiz, F., de Guzmán, I. G.-R., & Piattini, M. (2014). An integrated approach based on execution measures for the continuous improvement of business processes realized by services. *Information and Software Technology, 56*(2), 134–162. doi:10.1016/j.infsof.2013.08.003.

Delias, P., Doumpos, M., Grigoroudis, E., Manolitzas, P., & Matsatsinis, N. (2015). Supporting healthcare management decisions via robust clustering of event logs. *Knowledge-Based Systems, 84*, 203–213. doi:10.1016/j.knosys.2015.04.012.

Despaux, F., Song, Y.-Q., & Lahmadi, A. (2014). Towards performance analysis of wireless sensor networks using process mining techniques. *2014 IEEE symposium on computers and communications (ISCC)*. IEEE. doi:10.1109/iscc.2014.6912522.

Despaux, F., Song, Y.-Q., & Lahmadi, A. (2015). Extracting markov chain models from protocol execution traces for end to end delay evaluation in wireless sensor networks. *2015 IEEE world conference on factory communication systems (WFCS)*. IEEE. doi:10.1109/wfcs.2015.7160562.

Diamantini, C., Genga, L., Marozzo, F., Potena, D., & Trunfio, P. (2017). Discovering mobility patterns of instagram users through process mining techniques. *2017 IEEE international conference on information reuse and integration (IRI)*. IEEE. doi:10.1109/iri.2017.69.

Diamantini, C., Genga, L., Potena, D., & Storti, E. (2013). Pattern discovery from innovation processes. In *2013 international conference on collaboration technologies and systems (cts)* (pp. 457–464). doi:10.1109/CTS.2013.6567269.

Dijkman, R., Dumas, M., van Dongen, B., Käärik, R., & Mendling, J. (2011). Similarity of business process models: Metrics and evaluation. *Information Systems, 36*(2), 498–516. doi:10.1016/j.is.2010.09.006.

Dohrmann, P. (2014). Language-based process model discovery and enhancement. *2014 40th EUROMICRO conference on software engineering and advanced applications*. IEEE. doi:10.1109/seaa.2014.27.

Doleck, T., Jarrell, A., Poitras, E. G., Chaouachi, M., & Lajoie, S. P. (2016). Examining diagnosis paths: A process mining approach. In *2016 second international conference on computational intelligence communication technology (cict)* (pp. 663–667). doi:10.1109/CICT.2016.137.

van Dongen, B. (2007). An iterative algorithm for applying the theory of regions in process mining. *BETA publicaties: Preprints*. Beta, Research School for Operations Management and Logistics.

van Dongen, B., Jansen-Vullers, M., Verbeek, H., & van der Aalst, W. (2007). Verification of the SAP reference models using EPC reduction, state-space analysis, and invariants. *Computers in Industry, 58*(6), 578–601. doi:10.1016/j.compind.2007.01.001.

van Dongen, B. F., de Medeiros, A. K. A., Verbeek, H. M. W., Weijters, A. J. M. M., & van der Aalst, W. M. P. (2005). The ProM framework: A new era in process mining tool support. In *Applications and theory of petri nets 2005* (pp. 444–454). Springer Berlin Heidelberg. doi:10.1007/11494744_25.

Du, J., Cai, H., Jiang, L., & Huang, C. (2017). Methods of introducing continuous process mining to service management for mobile APPs. *2017 IEEE 14th international conference on e-business engineering (ICEBE)*. IEEE. doi:10.1109/icebe.2017.29.

Duan, B., & Shen, B. (2011). Software process discovery using link analysis. *2011 IEEE 3rd international conference on communication software and networks*. IEEE. doi:10.1109/iccsn.2011.6014218.

Dustdar, S., Hoffmann, T., & van der Aalst, W. (2005). Mining of ad-hoc business processes with TeamLog. *Data & Knowledge Engineering, 55*(2), 129–158. doi:10.1016/j.datak.2005.02.002.

van Eck, M. L., Sidorova, N., & van der Aalst, W. M. (2017). Guided interaction exploration in artifact-centric process models. *2017 IEEE 19th conference on business informatics (CBI)*. IEEE. doi:10.1109/cbi.2017.42.

Edgington, T. M., Raghu, T., & Vinze, A. S. (2010). Using process mining to identify coordination patterns in IT service management. *Decision Support Systems, 49*(2), 175–186. doi:10.1016/j.dss.2010.02.003.

Engel, R., & Bose, R. P. J. C. (2014). A case study on analyzing inter-organizational business processes from EDI messages using physical activity mining. *2014 47th hawaii international conference on system sciences*. IEEE. doi:10.1109/hicss.2014.479.

Erdogan, T., & Tarhan, A. (2016). Process mining for healthcare process analytics. *2016 joint conference of the international workshop on software measurement and the international conference on software process and product measurement (IWSM-MENSURA)*. IEEE. doi:10.1109/iwsm-mensura.2016.027.

Estrada-Vargas, A. P., Lopez-Mellado, E., & Lesage, J.-J. (2013). Automated modelling of reactive discrete event systems from external behavioural data. *CONIELECOMP 2013, 23rd international conference on electronics, communications and computing*. IEEE. doi:10.1109/conielecomp.2013.6525771.

Evermann, J., Rehse, J.-R., & Fettke, P. (2017). Predicting process behaviour using deep learning. *Decision Support Systems, 100*, 129–140. doi:10.1016/j.dss.2017.04.003. Smart Business Process Management

Fahland, D., Lo, D., & Maoz, S. (2013). Mining branching-time scenarios. *2013 28th IEEE/ACM international conference on automated software engineering (ASE)*. IEEE. doi:10.1109/ase.2013.6693102.

Fan, S., Li, X., & Zhao, J. L. (2012). Collaboration process patterns and efficiency of issue resolution in software development. *2012 international conference on collaboration technologies and systems (CTS)*. IEEE. doi:10.1109/cts.2012.6261105.

Fei, H., & Meskens, N. (2012). Clustering of patients' trajectories with an auto-stopped bisecting k-medoids algorithm. *Journal of Mathematical Modelling and Algorithms in Operations Research, 12*(2), 135–154. doi:10.1007/s10852-012-9198-0.

Ferdinand, P., & Emmanuel, H. (2015). From the audit trails to xes event logs facilitating process mining. *Studies in Health Technology and Informatics, 210*(Digital Healthcare Empowering Europeans), 40–44. doi:10.3233/978-1-61499-512-8-40.

Fernández-Gallego, B., Lama, M., Vidal, J. C., & Mucientes, M. (2013). Learning analytics framework for educational virtual worlds. *Procedia Computer Science, 25*, 443–447. doi:10.1016/j.procs.2013.11.056. 2013 International Conference on Virtual and Augmented Reality in Education.

Fernandez-Llatas, C., Bayo, J. L., Martinez-Romero, A., Benedi, J. M., & Traver, V. (2016). Interactive pattern recognition in cardiovascular disease management. a process mining approach. *2016 IEEE-EMBS international conference on biomedical and health informatics (BHI)*. IEEE. doi:10.1109/bhi.2016.7455906.

Fernandez-Llatas, C., Garcia-Gomez, J. M., Vicente, J., Naranjo, J. C., Robles, M., Benedi, J. M., et al. (2011). Behaviour patterns detection for persuasive design in nursing homes to help dementia patients. *2011 annual international conference of the IEEE engineering in medicine and biology society*. IEEE. doi:10.1109/iembs.2011.6091583.

Fernandez-Llatas, C., Martinez-Millana, A., Martinez-Romero, A., Benedi, J. M., & Traver, V. (2015). Diabetes care related process modelling using process mining techniques. lessons learned in the application of interactive pattern recognition: coping with the spaghetti effect. *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE. doi:10.1109/embc.2015.7318809.

Fernańdez-Llatas, C., Meneu, T., Benedi, J. M., & Traver, V. (2010). Activity-based process mining for clinical pathways computer aided design. *2010 annual international conference of the IEEE engineering in medicine and biology*. IEEE. doi:10.1109/iembs.2010.5627760.

Fernandez-Llatas, C., Sacchi, L., Benedi, J. M., Dagliati, A., Traver, V., & Bellazzi, R. (2014). Temporal abstractions to enrich activity-based process mining corpus with clinical time series. *IEEE-EMBS international conference on biomedical and health informatics (BHI)*. IEEE. doi:10.1109/bhi.2014.6864481.

Fernandez-Llatas, C., Valdivieso, B., Traver, V., & Benedi, J. M. (2014). Using process mining for automatic support of clinical pathways design. In *Methods in molecular biology* (pp. 79–88). Springer New York. doi:10.1007/978-1-4939-1985-7_5.

Ferreira, D., Zacarias, M., Malheiros, M., & Ferreira, P. (2007). Approaching process mining with sequence clustering: Experiments and findings. In *Lecture notes in computer science* (pp. 360–374). Springer Berlin Heidelberg. doi:10.1007/978-3-540-75183-0_26.

Flath, C. M., & Stein, N. (2018). Towards a data science toolbox for industrial analytics applications. *Computers in Industry, 94*, 16–25. doi:10.1016/j.compind.2017.09.003.

Forsberg, D., Rosipko, B., & Sunshine, J. L. (2015). Analyzing PACS usage patterns by means of process mining: Steps toward a more detailed workflow analysis in radiology. *Journal of Digital Imaging, 29*(1), 47–58. doi:10.1007/s10278-015-9824-2.

Francescomarino, C. D., Dumas, M., Maggi, F. M., & Teinemaa, I. (2017). Clustering-based predictive process monitoring. *IEEE Transactions on Services Computing*. doi:10.1109/tsc.2016.2645153. 1–1.

Fu, X., & Peng, Y. (2008). Automated capture of design knowledge from design process. *2008 fifth international conference on fuzzy systems and knowledge discovery*. IEEE. doi:10.1109/fskd.2008.223.

Funkner, A. A., Yakovlev, A. N., & Kovalchuk, S. V. (2017a). Data-driven modeling of clinical pathways using electronic health records. *Procedia Computer Science, 121*, 835–842. doi:10.1016/j.procs.2017.11.108.

Funkner, A. A., Yakovlev, A. N., & Kovalchuk, S. V. (2017b). Towards evolutionary discovery of typical clinical pathways in electronic health records. *Procedia Computer Science, 119*, 234–244. doi:10.1016/j.procs.2017.11.181.

Gadler, D., Mairegger, M., Janes, A., & Russo, B. (2017). Mining logs to model the use of a system. *2017 ACM/IEEE international symposium on empirical software engineering and measurement (ESEM)*. IEEE. doi:10.1109/esem.2017.47.

Ganesha, K., Dhanush, S., & Raj, S. M. S. (2017). An approach to fuzzy process mining to reduce patient waiting time in a hospital. *2017 international conference on innovations in information, embedded and communication systems (ICIIECS)*. IEEE. doi:10.1109/iciiecs.2017.8275889.

Ganesha, K., Raj, S. M. S., & Dhanush, S. (2017). Process mining approach for efficient utilization of resources in a hospital. *2017 international conference on innovations in information, embedded and communication systems (ICIIECS)*. IEEE. doi:10.1109/iciiecs.2017.8275886.

Ganesha, K., Soundarya, M., & Supriya, K. V. (2017). The best fit process model for the utilization of the physical resources in hospitals by applying inductive visual miner. *2017 international conference on inventive communication and computational technologies (ICICCT)*. IEEE. doi:10.1109/icicct.2017.7975212.

Ganesha, K., Supriya, K. V., & Soundarya, M. (2017). Analyzing the waiting time of patients in hospital by applying heuristics process miner. *2017 international conference on inventive communication and computational technologies (ICICCT)*. IEEE. doi:10.1109/icicct.2017.7975250.

Garcia, A. O., Ramirez, Y. E. P., & Larrea, O. U. A. (2015). Process mining in healthcare: Analysis and modeling of processes in the emergency area. *IEEE Latin America Transactions, 13*(5), 1612–1618. doi:10.1109/TLA.2015.7112022.

Garg, N., & Agarwal, S. (2016). Process mining for clinical workflows. In *Proceedings of the international conference on advances in information communication technology & computing - AICTC '16*. ACM Press. doi:10.1145/2979779.2979784.

Gatta, R., Montesi, M., Marchetti, A., Castellano, M., Valentini, V., Vallati, M., et al. (2017). Generating and comparing knowledge graphs of medical processes using pMineR. In *Proceedings of the knowledge capture conference on - k-CAP 2017*. ACM Press. doi:10.1145/3148011.3154464.

Gellrich, A., Wagner, T., Vasyutynskyy, V., & Kabitzsch, K. (2011). Modeling of transport times in partly observable factory logistic systems based on event logs. *ETFA2011*. IEEE. doi:10.1109/etfa.2011.6058992.

Genga, L. (2013). Application of process mining techniques for innovation analysis and support. In *2013 international conference on collaboration technologies and systems (cts)* (pp. 584–587). doi:10.1109/CTS.2013.6567291.

van Genuchten, M., Mans, R., Reijers, H., & Wismeijer, D. (2014). Is your upgrade worth it? Process mining can tell. *IEEE Software, 31*(5), 94–100. doi:10.1109/ms.2014.20.

Gerke, K. (2008). Impact of data aggregation on the significance of process mining results: An experimental evaluation. *2008 IEEE 17th workshop on enabling technologies: Infrastructure for collaborative enterprises*. IEEE. doi:10.1109/wetice.2008.25.

Gerke, K., Claus, A., & Mendling, J. (2009). Process mining of RFID-based supply chains. *2009 IEEE conference on commerce and enterprise computing*. IEEE. doi:10.1109/cec.2009.72.

Ghasemi, M., & Amyot, D. (2016). Process mining in healthcare: A systematised literature review. *International Journal of Electronic Healthcare, 9*(1), 60. doi:10.1504/ijeh.2016.078745.

Ghattas, J., Soffer, P., & Peleg, M. (2014). Improving business process decision making based on past experience. *Decision Support Systems, 59*, 93–107. doi:10.1016/j.dss.2013.10.009.

Giacalone, M., Cusatelli, C., & Santarcangelo, V. (2018). Big data compliance for innovative clinical models. *Big Data Research, 12*, 35–40. doi:10.1016/j.bdr.2018.02.001. Big Data Centric Computational Intelligence in Bioinformatics and Healthcare

Gillblad, D., Steinert, R., & Ferreira, D. R. (2009). Estimating the parameters of randomly interleaved markov models. *2009 IEEE international conference on data mining workshops*. IEEE. doi:10.1109/icdmw.2009.17.

Goedertier, S., Weerdt, J. D., Martens, D., Vanthienen, J., & Baesens, B. (2011). Process discovery in event logs: An application in the telecom industry. *Applied Soft Computing, 11*(2), 1697–1710. doi:10.1016/j.asoc.2010.04.025.

Gönczy, L., Heckel, R., & Varró, D. (2007). Model-based testing of service infrastructure components. In A. Petrenko, M. Veanes, J. Tretmans, & W. Grieskamp (Eds.), *Testing of software and communicating systems* (pp. 155–170). Berlin, Heidelberg: Springer Berlin Heidelberg.

Gonzalez, H., Han, J., & Li, X. (2006). Mining compressed commodity workflows from massive RFID data sets. In *Proceedings of the 15th ACM international conference on information and knowledge management - CIKM '06*. ACM Press. doi:10.1145/1183614.1183641.

Gou, Z., Tsugawa, S., Korkaew, A., Yamaguchi, S., Khamket, T., Manaskasemsak, B., et al. (2016). An interest-based tour planning tool by process mining from twitter. *2016 IEEE 5th global conference on consumer electronics*. IEEE. doi:10.1109/gcce.2016.7800542.

Greco, G., Guzzo, A., Pontieri, L., & Sacca, D. (2006). Discovering expressive process models by clustering log traces. *IEEE Transactions on Knowledge and Data Engineering, 18*(8), 1010–1027. doi:10.1109/tkde.2006.123.

Groba, A. R., Barreiros, B. V., Lama, M., Gewerc, A., & Mucientes, M. (2014). Using a learning analytics tool for evaluation in self-regulated learning. *2014 IEEE frontiers in education conference (FIE) proceedings*. IEEE. doi:10.1109/fie.2014.7044400.

Guarda, T., Santos, M. F., Augusto, M. F., Silva, C., & Pinto, F. (2013). Process mining: A framework proposal for pervasive business intelligence. In *2013 8th iberian conference on information systems and technologies (cisti)* (pp. 1–4).

Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy mining – Adaptive process simplification based on multi-perspective metrics. In *Lecture notes in computer science* (pp. 328–343). Springer Berlin Heidelberg. doi:10.1007/978-3-540-75183-0_24.

Guo, H., Brown, R., & Rasmussen, R. (2013). Workflow resource pattern modelling and visualization. In *Proceedings of the thirty-sixth australasian computer science conference - volume 135*. In *ACSC '13* (pp. 31–40). Darlinghurst, Australia, Australia: Australian Computer Society, Inc..

Gupta, M. (2014). Nirikshan: process mining software repositories to identify inefficiencies, imperfections, and enhance existing process capabilities. *Companion proceedings of the 36th international conference on software engineering - ICSE companion 2014*. ACM Press. doi:10.1145/2591062.2591080.

Gupta, M. (2017). Improving software maintenance using process mining and predictive analytics. *2017 IEEE international conference on software maintenance and evolution (ICSME)*. IEEE. doi:10.1109/icsme.2017.39.

Gupta, M., Asadullah, A., Padmanabhuni, S., & Serebrenik, A. (2017). Reducing user input requests to improve IT support ticket resolution process. *Empirical Software Engineering, 23*(3), 1664–1703. doi:10.1007/s10664-017-9532-2.

Gupta, M., & Sureka, A. (2014a). Nirikshan: Mining bug report history for discovering process maps, inefficiencies and inconsistencies. In *Proceedings of the 7th india software engineering conference*. In *ISEC '14* (pp. 1:1–1:10). New York, NY, USA: ACM. doi:10.1145/2590748.2590749.

Gupta, M., & Sureka, A. (2014b). Process cube for software defect resolution. *2014 21st Asia-Pacific software engineering conference*. IEEE. doi:10.1109/apsec.2014.45.

Gupta, M., Sureka, A., & Padmanabhuni, S. (2014). Process mining multiple repositories for software defect resolution from control and organizational perspective. In *Proceedings of the 11th working conference on mining software repositories - MSR 2014*. ACM Press. doi:10.1145/2597073.2597081.

Haigh, K. Z., & Yaman, F. (2011). RECYCLE. *ACM Transactions on Intelligent Systems and Technology, 2*(4), 1–32. doi:10.1145/1989734.1989746.

Halioui, A., Martin, T., Valtchev, P., & Diallo, A. B. (2017). Ontology-based workflow pattern mining. In *Proceedings of the symposium on applied computing - SAC '17*. ACM Press. doi:10.1145/3019612.3019866.

Han, B., Jiang, L., & Cai, H. (2011). Abnormal process instances identification method in healthcare environment. *2011 IEEE 10th international conference on trust, security and privacy in computing and communications*. IEEE. doi:10.1109/trustcom.2011.189.

Haung, K.-Z., Chen, Y.-S., & Chung, Y.-K. (2011). Perspectives on process mining within cloud computing. *2011 3rd international conference on advanced computer control*. IEEE. doi:10.1109/icacc.2011.6016496.

Haung, Z.-C., Chen, Y.-S., & Chung, Y.-K. (2013). Statistical process monitoring by using process mining. *IEEE conference anthology*. IEEE. doi:10.1109/anthology.2013.6785062.

He, X., Guo, J., Wang, Y., & Guo, Y. (2009). An automatic compliance checking approach for software processes. *2009 16th Asia-Pacific software engineering conference*. IEEE. doi:10.1109/apsec.2009.48.

Herbst, J., & Karagiannis, D. (2004). Workflow mining with InWoLvE. *Computers in Industry, 53*(3), 245–264. doi:10.1016/j.compind.2003.10.002.

Ho, G., Lau, H., Lee, C., Ip, A., & Pun, K. (2005). An intelligent production workflow mining system for continual quality enhancement. *The International Journal of Advanced Manufacturing Technology, 28*(7–8), 792–809. doi:10.1007/s00170-004-2416-9.

Ho, G. T. S., & Lau, H. C. W. (2007). Development of an OLAP-fuzzy based process mining system for quality improvement. In *IFIP international federation for information processing* (pp. 243–258). Springer US. doi:10.1007/978-0-387-44641-7_26.

Homayounfar, P. (2012). Process mining challenges in hospital information systems. In *2012 federated conference on computer science and information systems (fedcsis)* (pp. 1135–1140).

Hompes, B., Buijs, J., van der Aalst, W., Dixit, P., & Buurman, J. (2015). Discovering deviating cases and process variants using trace clustering. *27th benelux conference on artificial intelligence, 5–6 november 2015, hasselt, belgium*.

Horita, H., Hirayama, H., Tahara, Y., & Ohsuga, A. (2016). Goal achievement analysis based on LTL checking and decision tree for improvements of PAIS. In *Proceedings of the 31st annual ACM symposium on applied computing - SAC '16*. ACM Press. doi:10.1145/2851613.2851898.

Hsu, P.-Y., Chuang, Y.-C., Lo, Y.-C., & He, S.-C. (2017). Using contextualized activity-level duration to discover irregular process instances in business operations. *Information Sciences, 391–392*, 80–98. doi:10.1016/j.ins.2016.10.027.

Hu, W., Al-Dabbagh, A. W., Chen, T., & Shah, S. L. (2016). Process discovery of operator actions in response to univariate alarms. *IFAC-PapersOnLine, 49*(7), 1026–1031. doi:10.1016/j.ifacol.2016.07.337.

Huang, B., Li, C., Yin, C., & Zhao, X. (2012). Cloud manufacturing service platform for small- and medium-sized enterprises. *The International Journal of Advanced Manufacturing Technology, 65*(9–12), 1261–1272. doi:10.1007/s00170-012-4255-4.

Huang, B., Zhu, P., & Wu, C. (2012). Customer-centered careflow modeling based on guidelines. *Journal of Medical Systems, 36*(5), 3307–3319. doi:10.1007/s10916-012-9823-5.

Huang, C.-H., Lien, H.-L., & Wang, L. S.-L. (2016). An empirical case study of internet usage on student performance based on fuzzy association rules. In *Proceedings of the 3rd multidisciplinary international social networks conference on SocialInformatics 2016, data science 2016 - MISNC, SI, DS 2016*. ACM Press. doi:10.1145/2955129.2955132.

Huang, Z., Dong, W., Bath, P., Ji, L., & Duan, H. (2014). On mining latent treatment patterns from electronic medical records. *Data Mining and Knowledge Discovery, 29*(4), 914–949. doi:10.1007/s10618-014-0381-y.

Huang, Z., Dong, W., Duan, H., & Li, H. (2014). Similarity measure between patient traces for clinical pathway analysis: Problem, method, and applications. *IEEE Journal of Biomedical and Health Informatics, 18*(1), 4–14. doi:10.1109/jbhi.2013.2274281.

Huang, Z., Dong, W., Ji, L., Gan, C., Lu, X., & Duan, H. (2014). Discovery of clinical pathway patterns from event logs using probabilistic topic models. *Journal of Biomedical Informatics, 47*, 39–57. doi:10.1016/j.jbi.2013.09.003.

Huang, Z., Dong, W., Ji, L., He, C., & Duan, H. (2016). Incorporating comorbidities into latent treatment pattern mining for clinical pathways. *Journal of Biomedical Informatics, 59*, 227–239. doi:10.1016/j.jbi.2015.12.012.

Huang, Z., Dong, W., Ji, L., Yin, L., & Duan, H. (2015). On local anomaly detection and analysis for clinical pathways. *Artificial Intelligence in Medicine, 65*(3), 167–177. doi:10.1016/j.artmed.2015.09.001.

Huang, Z., Lu, X., & Duan, H. (2011). Mining association rules to support resource allocation in business process management. *Expert Systems with Applications, 38*(8), 9483–9490. doi:10.1016/j.eswa.2011.01.146.

Huang, Z., Lu, X., & Duan, H. (2012c). On mining clinical pathway patterns from medical behaviors. *Artificial Intelligence in Medicine, 56*(1), 35–50. doi:10.1016/j.artmed.2012.06.002.

Huang, Z., Lu, X., & Duan, H. (2013). Latent treatment pattern discovery for clinical processes. *Journal of Medical Systems, 37*(2). doi:10.1007/s10916-012-9915-2.

Huang, Z., Lu, X., Duan, H., & Fan, W. (2013). Summarizing clinical pathways from event logs. *Journal of Biomedical Informatics, 46*(1), 111–127. doi:10.1016/j.jbi.2012.10.001.

Huda, S., Ahmad, T., Sarno, R., & Santoso, H. A. (2014). Identification of process-based fraud patterns in credit application. *2014 2nd international conference on information and communication technology (ICoICT)*. IEEE. doi:10.1109/icoict.2014.6914045.

Hug, C., Deneckere, R., & Salinesi, C. (2012). Map-TBS: Map process enactment traces and analysis. *2012 sixth international conference on research challenges in information science (RCIS)*. IEEE. doi:10.1109/rcis.2012.6240435.

Huo, M., Zhang, H., & Jeffery, R. (2006a). An exploratory study of process enactment as input to software process improvement. In *Proceedings of the 2006 international workshop on software quality - WoSQ '06*. ACM Press. doi:10.1145/1137702.1137711.

Huo, M., Zhang, H., & Jeffery, R. (2006b). A systematic approach to process enactment analysis as input to software process improvement or tailoring. *2006 13th asia pacific software engineering conference (APSEC'06)*. IEEE. doi:10.1109/apsec.2006.14.

Ingvaldsen, J. E., & Gulla, J. A. (2012). Industrial application of semantic process mining. *Enterprise Information Systems, 6*(2), 139–163. doi:10.1080/17517575.2011.593103.

Intayoad, W., Kamyod, C., & Temdee, P. (2018). Process mining application for discovering student learning paths. *2018 international conference on digital arts, media and technology (ICDAMT)*. IEEE. doi:10.1109/icdamt.2018.8376527.

Jaisook, P., & Premchaiswadi, W. (2015). Time performance analysis of medical treatment processes by using disco. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368480.

Jangvaha, K., Porouhan, P., Palangsantikul, P., & Premchaiswadi, W. (2017). Analysis of emergency room service using fuzzy process mining technique. *2017 15th international conference on ICT and knowledge engineering (ICT&KE)*. IEEE. doi:10.1109/ictke.2017.8259634.

Jans, M., van der Werf, J. M., Lybaert, N., & Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications, 38*(10), 13351–13359. doi:10.1016/j.eswa.2011.04.159.

Jansen-Vullers, M., van der Aalst, W., & Rosemann, M. (2006). Mining configurable enterprise information systems. *Data & Knowledge Engineering, 56*(3), 195–244. doi:10.1016/j.datak.2005.03.007.

Janssenswillen, G., Depaire, B., & Verboven, S. (2017). Detecting train reroutings with process mining. *EURO Journal on Transportation and Logistics, 7*(1), 1–24. doi:10.1007/s13676-017-0105-8.

Jaroenphol, E., Porouhan, P., & Premchaiswadi, W. (2015). Analysis of the patients' treatment process in a hospital in thailand using fuzzy mining algorithms. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368484.

Jaturogpattana, T., Arpasat, P., Kungcharoen, K., Intarasema, S., & Premchaiswadi, W. (2017). Conformance analysis of outpatient data using process mining technique. *2017 15th international conference on ICT and knowledge engineering (ICT&KE)*. IEEE. doi:10.1109/ictke.2017.8259631.

Javed, M., Abgaz, Y. M., & Pahl, C. (2013). Ontology change management and identification of change patterns. *Journal on Data Semantics, 2*(2), 119–143. doi:10.1007/s13740-013-0024-2.

Jin, T., Wang, J., & Wen, L. (2007). Organizational modeling from event logs. *Sixth international conference on grid and cooperative computing (GCC 2007)*. IEEE. doi:10.1109/gcc.2007.93.

Jlailaty, D., Grigori, D., & Belhajjame, K. (2017). Business process instances discovery from email logs. *2017 IEEE international conference on services computing (SCC)*. IEEE. doi:10.1109/scc.2017.12.

Jo, H., Noh, S. D., & Cho, Y. (2014). An agile operations management system for green factory. *International Journal of Precision Engineering and Manufacturing-Green Technology, 1*(2), 131–143. doi:10.1007/s40684-014-0018-z.

Joe, J., Emmatty, T., Ballal, Y., & Kulkarni, S. (2016). Process mining for project management. *2016 international conference on data mining and advanced computing (SAPIENCE)*. IEEE. doi:10.1109/sapience.2016.7684142.

Joishi, J., & Sureka, A. (2014). Vishleshan. In *Proceedings of the 19th international database engineering & applications symposium on - IDEAS '15*. ACM Press. doi:10.1145/2790755.2790780.

Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior, 92*, 496–506. doi:10.1016/j.chb.2017.12.015.

Juneja, P., Kundra, D., & Sureka, A. (2016). Anvaya: An algorithm and case-study on improving the goodness of software process models generated by mining event-log data in issue tracking systems. *2016 IEEE 40th annual computer software and applications conference (COMPSAC)*. IEEE. doi:10.1109/compsac.2016.64.

Kanstren, T. (2009). Program comprehension for user-assisted test oracle generation. *2009 fourth international conference on software engineering advances*. IEEE. doi:10.1109/icsea.2009.26.

Karnok, D., & Monostori, L. (2011). Determination of routings and process time information from event logs. *IFAC Proceedings Volumes, 44*(1), 14055–14060. doi:10.3182/20110828-6-it-1002.01448.

Karray, M.-H., Chebel-Morello, B., & Zerhouni, N. (2014). PETRA: Process evolution using a TRAce-based system on a maintenance platform. *Knowledge-Based Systems, 68*, 21–39. doi:10.1016/j.knosys.2014.03.010.

Kaymak, U., Mans, R., van de Steeg, T., & Dierks, M. (2012). On process mining in health care. *2012 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE. doi:10.1109/icsmc.2012.6378009.

Kecman, P., & Goverde, R. M. (2013). Adaptive, data-driven, online prediction of train event times. *16th international IEEE conference on intelligent transportation systems (ITSC 2013)*. IEEE. doi:10.1109/itsc.2013.6728330.

Keertipati, S., Licorish, S. A., & Savarimuthu, B. T. R. (2016). Exploring decision-making processes in python. In *Proceedings of the 20th international conference on evaluation and assessment in software engineering - EASE '16*. ACM Press. doi:10.1145/2915970.2915993.

Kelleher, D. C., Bose, J. C., Waterhouse, L. J., Carter, E. A., & Burd, R. S. (2013). Checklist implementation reduces advanced trauma life support workflow deviations during trauma resuscitations without pre-arrival notification. *Journal of the American College of Surgeons, 217*(3), S101. doi:10.1016/j.jamcollsurg.2013.07.231.

Kelleher, D. C., Bose, R. J. C., Waterhouse, L. J., Carter, E. A., & Burd, R. S. (2014). Effect of a checklist on advanced trauma life support workflow deviations during trauma resuscitations without pre-arrival notification. *Journal of the American College of Surgeons, 218*(3), 459–466. doi:10.1016/j.jamcollsurg.2013.11.021.

Keski-Valkama, T. (2017). A simulator for event-oriented data in flexible assembly system fault prediction. *Procedia Computer Science, 119*, 121–130. doi:10.1016/j.procs.2017.11.168.

Khodabandelou, G., Hug, C., Deneckère, R., & Salinesi, C. (2014). Unsupervised discovery of intentional process models from event logs. In *Proceedings of the 11th working conference on mining software repositories - MSR 2014*. ACM Press. doi:10.1145/2597073.2597101.

Khodabandelou, G., Hug, C., & Salinesi, C. (2014). A novel approach to process mining: Intentional process models discovery. *2014 IEEE eighth international conference on research challenges in information science (RCIS)*. IEEE. doi:10.1109/rcis.2014.6861040.

Kim, E., Kim, S., Song, M., Kim, S., Yoo, D., Hwang, H., et al. (2013). Discovery of outpatient care process of a tertiary university hospital using process mining. *Healthcare Informatics Research, 19*(1), 42. doi:10.4258/hir.2013.19.1.42.

Kim, S., & Lee, K. (2013). Design of the integrated monitoring framework based on ontology for analyzing the customer feedback. *2013 international conference on information science and applications (ICISA)*. IEEE. doi:10.1109/icisa.2013.6579438.

Kingsley, O., Tawil, A. R. H., Naeem, U., Islam, S., & Lamine, E. (2016). Using semantic-based approach to manage perspectives of process mining: Application on improving learning process domain data. *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016*, 3529–3538. doi:10.1109/BigData.2016.7841016.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering.

Klausner, S., Entacher, K., Kranzer, S., Flamm, M., Sonnichsen, A., & Fritsch, G. (2014). ProPath - a guideline based software for the implementation into the medical environment. *2014 IEEE canada international humanitarian technology conference - (IHTC)*. IEEE. doi:10.1109/ihtc.2014.7147551.

Ko, R. K. L. (2009). A computer scientist's introductory guide to business process management (BPM). *Crossroads, 15*(4), 11–18. doi:10.1145/1558897.1558901.

Kondo, R. E., de F. R. Loures, E., & Santos, E. A. P. (2012). Process mining for alarm rationalization and fault patterns identification. In *Proceedings of 2012 IEEE 17th international conference on emerging technologies & factory automation (ETFA 2012)*. IEEE. doi:10.1109/etfa.2012.6489695.

Krathu, W., Pichler, C., Engel, R., Zapletal, M., Werthner, H., & Huemer, C. (2014). A framework for inter-organizational performance analysis from EDI messages. *2014 IEEE 16th conference on business informatics*. IEEE. doi:10.1109/cbi.2014.19.

Krathu, W., Pichler, C., Zapletal, M., & Werthner, H. (2012). Semantic inter-organizational performance analysis using the balanced scorecard methodology. In *2012 proceedings of the 35th international convention mipro* (pp. 1589–1594).

Krumeich, J., Jacobi, S., Werth, D., & Loos, P. (2014). Big data analytics for predictive manufacturing control - a case study from process industry. *2014 IEEE international congress on big data*. IEEE. doi:10.1109/bigdata.congress.2014.83.

Krutanard, C., Porouhan, P., & Premchaiswadi, W. (2015). Discovering organizational process models of resources in a hospital using role hierarchy miner. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368483.

Kudo, M., Nogayama, T., Ishida, A., & Abe, M. (2013). Business process analysis and real-world application scenarios. *2013 international conference on signal-image technology & internet-based systems*. IEEE. doi:10.1109/sitis.2013.159.

Kukreja, G., & Batra, S. (2017). Analogize process mining techniques in healthcare: Sepsis case study. In *2017 4th international conference on signal processing, computing and control (ispcc)* (pp. 482–487). doi:10.1109/ISPCC.2017.8269727.

Kumar, M. V. M., Thomas, L., & Annappa, B. (2017). Distilling lasagna from spaghetti processes. In *Proceedings of the 2017 international conference on intelligent systems, metaheuristics & swarm intelligence - ISMSI '17*. ACM Press. doi:10.1145/3059336.3059362.

Kurniati, A. P., Johnson, O., Hogg, D., & Hall, G. (2016). Process mining in oncology: A literature review. *2016 6th international conference on information communication and management (ICICM)*. IEEE. doi:10.1109/infocoman.2016.7784260.

Lübke, D. (2017). Extracting and conserving production data as test cases in executable business process architectures. *Procedia Computer Science, 121*, 1006–1013. doi:10.1016/j.procs.2017.11.130.

Lakshmanan, G. T., Mukhi, N. K., Khalaf, R., Martens, A., & Rozsnyai, S. (2012). Assessing the health of case-oriented semi-structured business processes. *2012 IEEE ninth international conference on services computing*. IEEE. doi:10.1109/scc.2012.92.

Lakshmanan, G. T., Shamsi, D., Doganata, Y. N., Unuvar, M., & Khalaf, R. (2013). A markov prediction model for data-driven semi-structured business processes. *Knowledge and Information Systems, 42*(1), 97–126. doi:10.1007/s10115-013-0697-8.

Lamma, E., Mello, P., Montali, M., Riguzzi, F., & Storari, S. (2007). Inducing declarative logic-based models from labeled traces. In G. Alonso, P. Dadam, & M. Rosemann (Eds.), *Business process management* (pp. 344–359). Berlin, Heidelberg: Springer Berlin Heidelberg.

Lau, H., Ho, G., Chu, K., Ho, W., & Lee, C. (2009a). Development of an intelligent quality management system using fuzzy association rules. *Expert Systems with Applications, 36*(2), 1801–1815. doi:10.1016/j.eswa.2007.12.066.

Lau, H., Ho, G., Zhao, Y., & Chung, N. (2009b). Development of a process mining system for supporting knowledge discovery in a supply chain network. *International Journal of Production Economics, 122*(1), 176–187. doi:10.1016/j.ijpe.2009.05.014.

kyung Lee, S., Kim, B., Huh, M., Cho, S., Park, S., & Lee, D. (2013). Mining transportation logs for understanding the after-assembly block manufacturing process in the shipbuilding industry. *Expert Systems with Applications, 40*(1), 83–95. doi:10.1016/j.eswa.2012.07.033.

Leemans, S. J. J., Fahland, D., & van der Aalst, W. M. P. (2013). Discovering block-structured process models from event logs - a constructive approach. In *Application and theory of petri nets and concurrency* (pp. 311–329). Springer Berlin Heidelberg. doi:10.1007/978-3-642-38697-8_17.

Leitner, M., & Rinderle-Ma, S. (2014). A systematic review on security in process-aware information systems – constitution, challenges, and future directions. *Information and Software Technology, 56*(3), 273–293. doi:10.1016/j.infsof.2013.12.004.

Lemos, A. M., Sabino, C. C., Lima, R. M. F., & Oliveira, C. A. L. (2011). Using process mining in software development process management: A case study. *2011 IEEE international conference on systems, man, and cybernetics*. IEEE. doi:10.1109/icsmc.2011.6083858.

de Leoni, M., & van der Aalst, W. M. P. (2013). Data-aware process mining. In *Proceedings of the 28th annual ACM symposium on applied computing - SAC '13*. ACM Press. doi:10.1145/2480362.2480633.

de Leoni, M., Maggi, F. M., & van der Aalst, W. M. (2015). An alignment-based framework to check the conformance of declarative process models and to preprocess event-log data. *Information Systems, 47*, 258–277. doi:10.1016/j.is.2013.12.005.

de Leoni, M., Suriadi, S., ter Hofstede, A. H. M., & van der Aalst, W. M. P. (2014). Turning event logs into process movies: animating what has really happened. *Software & Systems Modeling, 15*(3), 707–732. doi:10.1007/s10270-014-0432-2.

Letia, I. A., & Goron, A. (2015). Model checking as support for inspecting compliance to rules in flexible processes. *Journal of Visual Languages & Computing, 28*, 100–121. doi:10.1016/j.jvlc.2014.12.008.

Li, C., Reichert, M., & Wombacher, A. (2010). The minadept clustering approach for discovering reference process models out of process variants. *International Journal of Cooperative Information Systems, 19*(03n04), 159–203. doi:10.1142/s0218843010002139.

Li, C., Reichert, M., & Wombacher, A. (2011). Mining business process variants: Challenges, scenarios, algorithms. *Data & Knowledge Engineering, 70*(5), 409–434. doi:10.1016/j.datak.2011.01.005.

Li, H., Chan, K., Liang, M., & Luo, X. (2015). Composition of resource-service chain for cloud manufacturing. *IEEE Transactions on Industrial Informatics*. doi:10.1109/tii.2015.2503126. 1–1

Li, M., Liu, L., & Fan, B. (2008). A workflow-mining-based approach to knowledge maintenance. *2008 IEEE symposium on advanced management of information for globalized enterprises (AMIGE)*. IEEE. doi:10.1109/amige.2008.ecp.21.

Li, M., Liu, L., Yin, L., & Zhu, Y. (2010). A process mining based approach to knowledge maintenance. *Information Systems Frontiers, 13*(3), 371–380. doi:10.1007/s10796-010-9287-4.

Li, X.-Y., Yuan, J.-S., & Kong, Y.-H. (2007). Mining association rules from XML data with index table. *2007 international conference on machine learning and cybernetics*. IEEE. doi:10.1109/icmlc.2007.4370828.

Li, Y. (2009). Automatic remodeling of logistic business process in e-logistic. *2009 international conference on management science and engineering*. IEEE. doi:10.1109/icmse.2009.5317561.

Li, Y. (2009). A remodeling method of automatic learning process based on lms in e-learning. In *2009 international conference on web information systems and mining* (pp. 565–569). doi:10.1109/WISM.2009.120.

Li, Y. (2010). Automatic modeling of collaboration-based organization structure in e-logistic. *2010 international conference on intelligent computation technology and automation*. IEEE. doi:10.1109/icicta.2010.295.

Li, Y., & Deng, S. (2009). Automatic rediscovery and remodeling of logistic process in e-business logistic management. *2009 international conference on intelligent human-machine systems and cybernetics*. IEEE. doi:10.1109/ihmsc.2009.35.

Liao, Y., Deschamps, F., de Freitas Rocha Loures, E., & Ramos, L. F. P. (2017). Past, present and future of industry 4.0 - A systematic literature review and research agenda proposal. *International Journal of Production Research, 55*(12), 3609–3629. doi:10.1080/00207543.2017.1308576.

Lismont, J., Janssens, A.-S., Odnoletkova, I., vanden Broucke, S., Caron, F., & Vanthienen, J. (2016). A guide for the application of analytics on healthcare processes: A dynamic view on patient pathways. *Computers in Biology and Medicine, 77*, 125–134. doi:10.1016/j.compbiomed.2016.08.007.

Liu, C., van Dongen, B., Assy, N., & van der Aalst, W. M. (2016). Component behavior discovery from software execution data. *2016 IEEE symposium series on computational intelligence (SSCI)*. IEEE. doi:10.1109/ssci.2016.7849947.

Liu, H., Liu, Y., Li, X., Xie, G., & Lakshmanan, G. T. (2014). Towards pathway variation identification. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management - CIKM '14*. ACM Press. doi:10.1145/2661829.2662013.

Liu, T., Yi, H., Ni, Z., & Liu, X. (2008). A closed-loop workflow management technique based on process mining. *2008 15th international conference on mechatronics and machine vision in practice*. IEEE. doi:10.1109/mmvip.2008.4749573.

Lopez, J., Maag, S., Saint-Pierre, C., Bustos, J., & Cavalli, A. (2015). Process mining for trust monitoring. *2015 IEEE 29th international conference on advanced information networking and applications workshops*. IEEE. doi:10.1109/waina.2015.71.

Lou, J.-G., Fu, Q., Yang, S., Li, J., & Wu, B. (2010). Mining program workflow from interleaved traces. In *Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining - KDD '10*. ACM Press. doi:10.1145/1835804.1835883.

Low, W., van der Aalst, W., ter Hofstede, A., Wynn, M., & Weerdt, J. D. (2017). Change visualisation: Analysing the resource and timing differences between two event logs. *Information Systems, 65*, 106–123. doi:10.1016/j.is.2016.10.005.

Lu, F., Zeng, Q., & Duan, H. (2016). Synchronization-core-based discovery of processes with decomposable cyclic dependencies. *ACM Transactions on Knowledge Discovery from Data, 10*(3), 1–29. doi:10.1145/2845086.

Ly, L. T., Maggi, F. M., Montali, M., Rinderle-Ma, S., & van der Aalst, W. M. (2013). A framework for the systematic comparison and evaluation of compliance monitoring approaches. *2013 17th IEEE international enterprise distributed object computing conference*. IEEE. doi:10.1109/edoc.2013.11.

Ly, L. T., Maggi, F. M., Montali, M., Rinderle-Ma, S., & van der Aalst, W. M. (2015). Compliance monitoring in business processes: Functionalities, application, and tool-support. *Information Systems, 54*, 209–234. doi:10.1016/j.is.2015.02.007.

Maaradji, A., Dumas, M., Rosa, M. L., & Ostovar, A. (2017). Detecting sudden and gradual drifts in business processes from execution traces. *IEEE Transactions on Knowledge and Data Engineering, 29*(10), 2140–2154. doi:10.1109/tkde.2017.2720601.

Ma'arif, M. R. (2017). Revealing daily human activity pattern using process mining approach. *2017 4th international conference on electrical engineering, computer science and informatics (EECSI)*. IEEE. doi:10.1109/eecsi.2017.8239160.

Mahendrawathi, E., Astuti, H. M., & Nastiti, A. (2015). Analysis of customer fulfilment with process mining: A case study in a telecommunication company. *Procedia Computer Science, 72*, 588–596. doi:10.1016/j.procs.2015.12.167.

Mahmood, T., & Shaikh, G. M. (2013). Adaptive automated teller machines. *Expert Systems with Applications, 40*(4), 1152–1169. doi:10.1016/j.eswa.2012.08.022.

Maita, A., Corrêa Martins, L., Paz, C., Rafferty, L., Hung, P., Peres, S., et al. (2018). A systematic mapping study of process mining. *Enterprise Information Systems, 12*, 505–549. doi:10.1080/17517575.2017.1402371.

Maita, A. R. C., Fantinato, M., Peres, S. M., Thom, L. H., & Hung, P. C. K. (2017). Mining unstructured processes: An exploratory study on a distance learning domain. In *2017 international joint conference on neural networks (ijcnn)* (pp. 3240–3247). doi:10.1109/IJCNN.2017.7966261.

Maldonado, J. J., Pérez-Sanagustín, M., Bermeo, J. L., Muñoz, L., Pacheco, G., & Espinoza, I. (2017). Flipping the classroom with moocs. a pilot study exploring differences between self-regulated learners. In *2017 twelfth latin american conference on learning technologies (laclo)* (pp. 1–8). doi:10.1109/LACLO.2017.8120934.

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses. *Computers in Human Behavior, 80*, 179–196. doi:10.1016/j.chb.2017.11.011.

Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high- and low-performing groups. *Computers in Human Behavior, 52*, 562–572. doi:10.1016/j.chb.2015.03.082.

Mans, R., Reijers, H., van Genuchten, M., & Wismeijer, D. (2012). Mining processes in dentistry. In *Proceedings of the 2nd ACM SIGHIT symposium on international health informatics - IHI '12*. ACM Press. doi:10.1145/2110363.2110407.

Mans, R., Reijers, H., Wismeijer, D., & van Genuchten, M. (2013). A process-oriented methodology for evaluating the impact of IT: A proposal and an application in healthcare. *Information Systems, 38*(8), 1097–1115. doi:10.1016/j.is.2013.06.005.

Mans, R. S., van der Aalst, W. M. P., & Vanwersch, R. J. B. (2015). *Process Mining in Healthcare*. Springer International Publishing. doi:10.1007/978-3-319-16071-9.

Mapikou, G. L. M., & Etoundi, R. A. (2016). A process mining oriented approach to improve process models analysis in developing countries. *2016 IEEE/ACS 13th international conference of computer systems and applications (AICCSA)*. IEEE. doi:10.1109/aiccsa.2016.7945759.

Markowski, P., & Przybyłek, M. (2017). Process mining methods for post-delivery validation. In *Proceedings of the 2017 federated conference on computer science and information systems*. IEEE. doi:10.15439/2017f372.

Márquez-Chamorro, A. E., Resinas, M., Ruiz-Cortés, A., & Toro, M. (2017). Run-time prediction of business process indicators using evolutionary decision rules. *Expert Systems with Applications, 87*, 1–14. doi:10.1016/j.eswa.2017.05.069.

Martinovic, J., Dráždilová, P., Slaninová, K., Kocyan, T., & Snasel, V. (2012). Left-right oscillate algorithm for community detection used in e-learning system. In *Computer information systems and industrial management* (pp. 278–289). doi:10.1007/978-3-642-33260-9_24.

Maruster, L., & Jorna, R. (2005). From data to knowledge: A method for modeling hospital logistic processes. *IEEE Transactions on Information Technology in Biomedicine, 9*(2), 248–255. doi:10.1109/titb.2005.847194.

Mastroianni, C., & Papuzzo, G. (2014). A self-organizing p2p framework for collective service discovery. *Journal of Network and Computer Applications, 39*, 214–222. doi:10.1016/j.jnca.2013.07.002.

de Medeiros, A. K. A., & van der Aalst, W. M. P. (2008). Process mining towards semantics. In *Advances in web semantics i* (pp. 35–80). Springer Berlin Heidelberg. doi:10.1007/978-3-540-89784-2_3.

Meethaisong, W., & Premchaiswadi, W. (2015). Applying social network miner on medical event logs using handover of work metric. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368481.

Meier, J., Dietz, A., Boehm, A., & Neumuth, T. (2015). Predicting treatment process steps from events. *Journal of Biomedical Informatics, 53*, 308–319. doi:10.1016/j.jbi.2014.12.003.

Meincheim, A., dos Santos Garcia, C., Nievola, J. C., & Scalabrin, E. E. (2017). Combining process mining with trace clustering: Manufacturing shop floor process - an applied case. *2017 IEEE 29th international conference on tools with artificial intelligence (ICTAI)*. IEEE. doi:10.1109/ictai.2017.00082.

Meng, F., Ooi, C. K., Soh, C. K. K., Teow, K. L., & Kannapiran, P. (2016). Quantifying patient flow and utilization with patient flow pathway and diagnosis of an emergency department in singapore. *Health Systems, 5*(2), 140–148. doi:10.1057/hs.2015.15.

Metsker, O., Bolgova, E., Yakovlev, A., Funkner, A., & Kovalchuk, S. (2017). Pattern-based mining in electronic health records for complex clinical process analysis. *Procedia Computer Science, 119*, 197–206. doi:10.1016/j.procs.2017.11.177.

Micieli, G., Cavallini, A., Quaglini, S., Fontana, G., & Duè, M. (2010). The lombardia stroke unit registry: 1-year experience of a web-based hospital stroke registry. *Neurological Sciences, 31*(5), 555–564. doi:10.1007/s10072-010-0249-x.

Misev, A., & Atanassov, E. (2009). ULMon - grid nonitoring from user point of view. In *Proceedings of the ITI 2009 31st international conference on information technology interfaces*. IEEE. doi:10.1109/iti.2009.5196158.

Mitsyuk, A. A. (2017). Non-local correction of process models using event logs. *2017 Ivannikov ISPRAS open conference (ISPRAS)*. IEEE. doi:10.1109/ispras.2017.00008.

Mittal, M., & Sureka, A. (2014). Process mining software repositories from student projects in an undergraduate software engineering course. *Companion proceedings of the 36th international conference on software engineering - ICSE companion 2014*. ACM Press. doi:10.1145/2591062.2591152.

Monroy, C., Rangel, V. S., Bell, E. R., & Whitaker, R. (2015). A learning analytics approach to characterize and analyze inquiry-based pedagogical processes. In *Proceedings of the fifth international conference on learning analytics and knowledge - LAK '15*. ACM Press. doi:10.1145/2723576.2723658.

Montani, S., Leonardi, G., Quaglini, S., Cavallini, A., & Micieli, G. (2014). Improving structural medical process comparison by exploiting domain knowledge and mined information. *Artificial Intelligence in Medicine, 62*(1), 33–45. doi:10.1016/j.artmed.2014.07.001.

Montani, S., Leonardi, G., Striani, M., Quaglini, S., & Cavallini, A. (2017). Multi-level abstraction for trace comparison and process discovery. *Expert Systems with Applications, 81*, 398–409. doi:10.1016/j.eswa.2017.03.063.

Mukala, P., Cerone, A., & Turini, F. (2017). An empirical verification of a-priori learning models on mailing archives in the context of online learning activities of participants in free\libre open source software (FLOSS) communities. *Education and Information Technologies, 22*(6), 3207–3229. doi:10.1007/s10639-017-9573-6.

Mulyar, N., van der Aalst, W. M. P., & Peleg, M. (2007). A pattern-based analysis of clinical computer-interpretable guideline modeling languages. *Journal of the American Medical Informatics Association, 14*(6), 781–787. doi:10.1197/jamia.m2389.

Mundbrod, N., Beuter, F., & Reichert, M. (2015). Supporting knowledge-intensive processes through integrated task lifecycle support. *2015 IEEE 19th international enterprise distributed object computing conference*. IEEE. doi:10.1109/edoc.2015.13.

Munoz-Gama, J., & Echizen, I. (2012). Insuring sensitive processes through process mining. *2012 9th international conference on ubiquitous intelligence and computing and 9th international conference on autonomic and trusted computing*. IEEE. doi:10.1109/uic-atc.2012.83.

Myers, D., Suriadi, S., Radke, K., & Foo, E. (2018). Anomaly detection for industrial control systems using process mining. *Computers & Security, 78*, 103–125. doi:10.1016/j.cose.2018.06.002.

Neamsirorat, W., & Premchaiswadi, W. (2015). Analysis of surgical event logs in a hospital by using heuristics miner technique. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368479.

Neyem, A., Diaz-Mosquera, J., Munoz-Gama, J., & Navon, J. (2017). Understanding student interactions in capstone courses to improve learning experiences. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education - SIGCSE '17*. ACM Press. doi:10.1145/3017680.3017716.

Ogasawara, S., Tayama, K., & Yamamura, T. (2006). Estimating structures of business process models from execution logs. *2006 10th IEEE international enterprise distributed object computing conference (EDOC'06)*. IEEE. doi:10.1109/edoc.2006.25.

Okoye, K., Tawil, A.-R. H., Naeem, U., Bashroush, R., & Lamine, E. (2014). A semantic rule-based approach supported by process mining for personalised adaptive learning. *Procedia Computer Science, 37*, 203–210. doi:10.1016/j.procs.2014.08.031.

Okoye, K., Tawil, A. R. H., Naeem, U., Bashroush, R., & Lamine, E. (2014). A semantic rule-based approach towards process mining for personalised adaptive learning. In *2014 ieee intl conf on high performance computing and communications, 2014 ieee 6th intl symp on cyberspace safety and security, 2014 IEEE 11th intl conf on embedded software and syst (hpcc,css,icess)* (pp. 929–936). doi:10.1109/HPCC.2014.143.

Ou-Yang, C., & Juan, Y.-C. (2010). Applying process mining approach to support the verification of a multi-agent system. *Journal of Systems Science and Systems Engineering, 19*(2), 131–149. doi:10.1007/s11518-010-5132-z.

Pane, E. S., Wibawa, A. D., & Purnomo, M. H. (2016). Event log-based fraud rating using interval type-2 fuzzy sets in fuzzy AHP. *2016 IEEE region 10 conference (TENCON)*. IEEE. doi:10.1109/tencon.2016.7848366.

Papadopoulos, A. N., Tsiouris, K. M., Pappas, I. G., Zervakis, M. E., Koutsouris, D. D., Exarchos, T. P., et al. (2015). Design and implementation of processes for the primary care in the healthcare system of greece. *2015 IEEE 15th international conference on bioinformatics and bioengineering (BIBE)*. IEEE. doi:10.1109/bibe.2015.7367717.

Park, J., Lee, D., & Zhu, J. (2014). An integrated approach for ship block manufacturing process performance evaluation: Case from a korean shipbuilding company. *International Journal of Production Economics, 156*, 214–222. doi:10.1016/j.ijpe.2014.06.012.

Park, S., & Kang, Y. S. (2016). A study of process mining-based business process innovation. *Procedia Computer Science, 91*, 734–743. doi:10.1016/j.procs.2016.07.066.

Partington, A., Wynn, M., Suriadi, S., Ouyang, C., & Karnon, J. (2015). Process mining for clinical processes. *ACM Transactions on Management Information Systems, 5*(4), 1–18. doi:10.1145/2629446.

Patel, P., Ali, M. I., & Sheth, A. (2018). From raw data to smart manufacturing: Ai and semantic web of things for industry 4.0. *IEEE Intelligent Systems, 33*(4), 79–86. doi:10.1109/MIS.2018.043741325.

Peleg, M. (2013). Computer-interpretable clinical guidelines: A methodological review. *Journal of Biomedical Informatics, 46*(4), 744–763. doi:10.1016/j.jbi.2013.06.009.

Pérez-Castillo, R., Cruz-Lemus, J. A., de Guzmán, I. G.-R., & Piattini, M. (2012). A family of case studies on business process mining using MARBLE. *Journal of Systems and Software, 85*(6), 1370–1385. doi:10.1016/j.jss.2012.01.022.

Pérez-Castillo, R., Weber, B., de Guzmán, I. G.-R., Piattini, M., & Pinggera, J. (2012a). Assessing event correlation in non-process-aware information systems. *Software & Systems Modeling, 13*(3), 1117–1139. doi:10.1007/s10270-012-0285-5.

Pérez-Castillo, R., Weber, B., de Guzmán, I. G.-R., Piattini, M., & Pinggera, J. (2012b). Assessing event correlation in non-process-aware information systems. *Software & Systems Modeling*. doi:10.1007/s10270-012-0285-5.

Peŕez-Castillo, R., Weber, B., de Guzmań, I.-R., & Piattini, M. (2011). Process mining through dynamic analysis for modernising legacy systems. *IET Software, 5*(3), 304. doi:10.1049/iet-sen.2010.0103.

Perimal-Lewis, L., De Vries, D., & Thompson, C. H. (2014). Health intelligence: Discovering the process model using process mining by constructing start-to-end patient journeys. In *Proceedings of the seventh australasian workshop on health informatics and knowledge management - volume 153*. In *HIKM '14* (pp. 59–67). Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

Perimal-Lewis, L., Qin, S., Thompson, C., & Hakendorf, P. (2012). Gaining insight from patient journey data using a process-oriented analysis approach. In *Proceedings of the fifth australasian workshop on health informatics and knowledge management - volume 129*. In *HIKM '12* (pp. 59–66). Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

Perimal-Lewis, L., Qin, S., Thompson, C. H., & Hakendorf, P. (2013). Analysing homogenous patient journeys to assess quality of care for patients admitted outside of their 'home-ward'. In *Proceedings of the sixth australasian workshop on health informatics and knowledge management - volume 142*. In *HIKM '13* (pp. 29–37). Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

Peters, E. M., Dedene, G., & Poelmans, J. (2013). Understanding service quality and customer churn by process discovery for a multi-national banking contact center. *2013 IEEE 13th international conference on data mining workshops*. IEEE. doi:10.1109/icdmw.2013.65.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. *Ease*.

Pika, A., van der Aalst, W., Wynn, M., Fidge, C., & ter Hofstede, A. (2016). Evaluating and predicting overall process risk using event logs. *Information Sciences, 352–353*, 98–120. doi:10.1016/j.ins.2016.03.003.

del Pilar Villamil, M., Barrera, D., Velasco, N., Bernal, O., Fajardo, E., Urango, C., et al. (2017). Strategies for the quality assessment of the health care service providers in the treatment of gastric cancer in colombia. *BMC Health Services Research, 17*(1). doi:10.1186/s12913-017-2440-8.

Pla, A., Gay, P., Meléndez, J., & López, B. (2012). Petri net-based process monitoring: a workflow management system for process modelling and monitoring. *Journal of Intelligent Manufacturing, 25*(3), 539–554. doi:10.1007/s10845-012-0704-z.

Polato, M., Sperduti, A., Burattin, A., & de Leoni, M. (2014). Data-aware remaining time prediction of business process instances. *2014 international joint conference on neural networks (IJCNN)*. IEEE. doi:10.1109/ijcnn.2014.6889360.

Poncin, W., Serebrenik, A., & van den Brand, M. (2011a). Mining student capstone projects with FRASR and ProM. In *Proceedings of the ACM international conference companion on object oriented programming systems languages and applications companion - SPLASH '11*. ACM Press. doi:10.1145/2048147.2048181.

Poncin, W., Serebrenik, A., & van den Brand, M. (2011b). Process mining software repositories. *2011 15th european conference on software maintenance and reengineering*. IEEE. doi:10.1109/csmr.2011.5.

Porouhan, P., Jongsawat, N., & Premchaiswadi, W. (2014). Process and deviation exploration through alpha-algorithm and heuristic miner techniques. *2014 twelfth international conference on ICT and knowledge engineering*. IEEE. doi:10.1109/ictke.2014.7001540.

Porouhan, P., & Premchaiswadi, W. (2015). Social network analysis and role hierarchy mining through MXML-based event logs. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368471.

Porouhan, P., & Premchaiswadi, W. (2016). Development of a process-aware instructor-aware multi-tabletop collaborative learning environment. In *2016 14th international conference on ICT and knowledge engineering (ict ke)* (pp. 62–70). doi:10.1109/ICTKE.2016.7804100.

Potavin, J., Jongswat, N., & Premchaiswadi, W. (2012). Applying fuzzy-genetic mining in conformance and dependency relations. In *2012 tenth international conference on ict and knowledge engineering* (pp. 228–235). doi:10.1109/ICTKE.2012.6408560.

Premchaiswadi, W., & Porouhan, P. (2015a). Process modeling and bottleneck mining in online peer-review systems. *SpringerPlus, 4*(1). doi:10.1186/s40064-015-1183-4.

Premchaiswadi, W., & Porouhan, P. (2015b). Process modeling and decision mining in a collaborative distance learning environment. *Decision Analytics, 2*(1), 6. doi:10.1186/s40165-015-0015-5.

Priyambada, S. A., Mahendrawathi, E., & Yahya, B. N. (2017). Curriculum assessment of higher educational institution using aggregate profile clustering. *Procedia Computer Science, 124*, 264–273. doi:10.1016/j.procs.2017.12.155. 4th Information Systems International Conference 2017, ISICO 2017, 6–8 November 2017, Bali, Indonesia

Prodel, M., Augusto, V., Jouaneton, B., Lamarsalle, L., & Xie, X. (2018). Optimal process mining for large and complex event logs. *IEEE Transactions on Automation Science and Engineering, 15*(3), 1309–1325. doi:10.1109/tase.2017.2784436.

Prodel, M., Augusto, V., Xie, X., Jouaneton, B., & Lamarsalle, L. (2015). Discovery of patient pathways from a national hospital database using process mining and integer linear programming. *2015 IEEE international conference on automation science and engineering (CASE)*. IEEE. doi:10.1109/coase.2015.7294295.

Prodel, M., Augusto, V., Xie, X., Jouaneton, B., & Lamarsalle, L. (2017). Stochastic simulation of clinical pathways from raw health databases. *2017 13th IEEE conference on automation science and engineering (CASE)*. IEEE. doi:10.1109/coase.2017.8256167.

Pulshashi, I. R., Bae, H., Sutrisnowati, R. A., Yahya, B. N., & Park, S. (2015). Slice and connect: Tri-dimensional process discovery with case study of port logistics process. *Procedia Computer Science, 72*, 461–468. doi:10.1016/j.procs.2015.12.127.

Qiao, C., & Hu, X. (2018). Discovering student behavior patterns from event logs: Preliminary results on a novel probabilistic latent variable model. *2018 IEEE 18th international conference on advanced learning technologies (ICALT)*. IEEE. doi:10.1109/icalt.2018.00056.

Rahmawati, D., & Sarno, R. (2017). Scalable attack analysis of business process based on decision mining classification. *2017 4th international conference on electrical engineering, computer science and informatics (EECSI)*. IEEE. doi:10.1109/eecsi.2017.8239135.

Rahmawati, D., Sarno, R., Fatichah, C., & Sunaryono, D. (2017). Fraud detection on event log of bank financial credit business process using hidden markov model algorithm. *2017 3rd international conference on science in information technology (ICSITech)*. IEEE. doi:10.1109/icsitech.2017.8257082.

Raichelson, L., Soffer, P., & Verbeek, E. (2017). Merging event logs: Combining granularity levels for process flow analysis. *Information Systems, 71*, 211–227. doi:10.1016/j.is.2017.08.010.

Rattanavayakorn, P., & Premchaiswadi, W. (2015). Analysis of the social network miner (working together) of physicians. *2015 13th international conference on ICT and knowledge engineering (ICT & knowledge engineering 2015)*. IEEE. doi:10.1109/ictke.2015.7368482.

R'bigui, H., & Cho, C. (2017). Customer order fulfillment process analysis with process mining. In *Proceedings of the 2017 international conference on computer science and artificial intelligence - CSAI 2017*. ACM Press. doi:10.1145/3168390.3168415.

Rebuge, Á., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. *Information Systems, 37*(2), 99–116. doi:10.1016/j.is.2011.01.003. Management and Engineering of Process-Aware Information Systems

Rebuge, A., Lapao, L. V., Freitas, A., & Cruz-Correia, R. (2013). A process mining analysis on a virtual electronic patient record system. In *Proceedings of the 26th IEEE international symposium on computer-based medical systems*. IEEE. doi:10.1109/cbms.2013.6627874.

Recker, J., & Mendling, J. (2015). The state of the art of business process management research as published in the BPM conference. *Business & Information Systems Engineering, 58*(1), 55–72. doi:10.1007/s12599-015-0411-3.

Reijers, H., Vanderfeesten, I., & van der Aalst, W. (2016). The effectiveness of workflow management systems: A longitudinal study. *International Journal of Information Management, 36*(1), 126–141. doi:10.1016/j.ijinfomgt.2015.08.003.

Reimann, P., Frerejean, J., & Thompson, K. (2009). Using process mining to identify models of group decision making in chat data. In *Proceedings of the 9th international conference on computer supported collaborative learning - volume 1*. In *CSCL'09* (pp. 98–107). International Society of the Learning Sciences.

Repta, D., & Stanescu, A. M. (2017). Holistic approach for process model discovery. *2017 21st international conference on control systems and computer science (CSCS)*. IEEE. doi:10.1109/cscs.2017.73.

Riaño, D., & Ortega, W. (2017). Computer technologies to integrate medical treatments to manage multimorbidity. *Journal of Biomedical Informatics, 75*, 1–13. doi:10.1016/j.jbi.2017.09.009.

Robinson, W. N., & Deng, T. (2015). Data mining behavioral transitions in open source repositories. *2015 48th Hawaii international conference on system sciences*. IEEE. doi:10.1109/hicss.2015.622.

Rodriguez, M. C., Nistal, M. L., Fonte, F. A. M., Penin, M. L., & Molina, M. M. (2018). Exploring the application of process mining to support self-regulated learning: An initial analysis with video lectures. *2018 IEEE global engineering education conference (EDUCON)*. IEEE. doi:10.1109/educon.2018.8363448.

Rogge-Solti, A., Mans, R. S., van der Aalst, W. M. P., & Weske, M. (2013). Improving documentation by repairing event logs. In *Lecture notes in business information processing* (pp. 129–144). Springer Berlin Heidelberg. doi:10.1007/978-3-642-41641-5_10.

Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. *Journal of Biomedical Informatics, 61*, 224–236. doi:10.1016/j.jbi.2016.04.007.

Rojas, E., Sepúlveda, M., Munoz-Gama, J., Capurro, D., Traver, V., & Fernandez-Llatas, C. (2017). Question-driven methodology for analyzing emergency room processes using process mining. *Applied Sciences, 7*(3), 302. doi:10.3390/app7030302.

Rosa, M. L., Aalst, W. M. P. V. D., Dumas, M., & Milani, F. P. (2017). Business process variability modeling: A survey. *ACM Comput. Surv., 50*(1), 2:1–2:45. doi:10.1145/3041957.

Rosa, N. (2017). Middleware adaptation through process mining. *2017 IEEE 31st international conference on advanced information networking and applications (AINA)*. IEEE. doi:10.1109/aina.2017.25.

Rosa, N., Campos, G., & Cavalcanti, D. (2017). Using software architecture principles and lightweight formalisation to build adaptive middleware. In *Proceedings of the 16th workshop on adaptive and reflective middleware - ARM '17*. ACM Press. doi:10.1145/3152881.3152882.

Rovani, M., Maggi, F. M., de Leoni, M., & van der Aalst, W. M. (2015). Declarative process mining in healthcare. *Expert Systems with Applications, 42*(23), 9236–9251. doi:10.1016/j.eswa.2015.07.040.

Rozinat, A., de Jong, I., Gunther, C., & van der Aalst, W. (2009). Process mining applied to the test process of wafer scanners in ASML. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 39*(4), 474–479. doi:10.1109/tsmcc.2009.2014169.

Rozinat, A., Mans, R., Song, M., & van der Aalst, W. (2009). Discovering simulation models. *Information Systems, 34*(3), 305–327. doi:10.1016/j.is.2008.09.002.

Rozinat, A., Wynn, M., van der Aalst, W., ter Hofstede, A., & Fidge, C. (2009). Workflow simulation for operational decision support. *Data & Knowledge Engineering, 68*(9), 834–850. doi:10.1016/j.datak.2009.02.014.

Rozsnyai, S., Lakshmanan, G. T., Muthusamy, V., Khalaf, R., & Duftler, M. J. (2012). Business process insight: An approach and platform for the discovery and analysis of end-to-end business processes. *2012 annual SRII global conference*. IEEE. doi:10.1109/srii.2012.20.

Rubin, V., Günther, C. W., van der Aalst, W. M. P., Kindler, E., van Dongen, B. F., & Schäfer, W. (2007). Process mining framework for software processes. In Q. Wang, D. Pfahl, & D. M. Raffo (Eds.), *Software process dynamics and agility* (pp. 169–181). Berlin, Heidelberg: Springer Berlin Heidelberg.

Rubin, V., Lomazova, I., & van der Aalst, W. M. P. (2014). Agile development with software process mining. In *Proceedings of the 2014 international conference on software and system process - ICSSP 2014*. ACM Press. doi:10.1145/2600821.2600842.

Ruschel, E., Santos, E. A. P., & de Freitas Rocha Loures, E. (2017). Mining shop-floor data for preventive maintenance management: Integrating probabilistic and predictive models. *Procedia Manufacturing, 11*, 1127–1134. doi:10.1016/j.promfg.2017.07.234.

Ruschel, E., Santos, E. A. P., & de Freitas Rocha Loures, E. (2018). Establishment of maintenance inspection intervals: an application of process mining techniques in manufacturing. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-018-1434-7.

Russell, N., van der Aalst, W. M. P., ter Hofstede, A. H. M., & Edmond, D. (2005). Workflow resource patterns: Identification, representation and tool support. In *Advanced information systems engineering, 17th international conference, Caise 2005, Porto, Portugal* (pp. 216–232). http://link.springer.com/chapter/10.1007/11431855_16. doi:10.1007/11431855_16.

Saelim, N., Porouhan, P., & Premchaiswadi, W. (2016). Improving organizational process of a hospital through petri-net based repair models. *2016 14th international conference on ICT and knowledge engineering (ICT&KE)*. IEEE. doi:10.1109/ictke.2016.7804107.

Sailer, A., Deubzer, M., Luttgen, G., & Mottok, J. (2016). CoreTAna: A trace analyzer for reverse engineering real-time software. *2016 IEEE 23rd international conference on software analysis, evolution, and reengineering (SANER)*. IEEE. doi:10.1109/saner.2016.36.

Saint-Pierre, C., Cifuentes, F., & Bustos-Jimenez, J. (2014). Detecting anomalies in DNS protocol traces via passive testing and process mining. *2014 IEEE conference on communications and network security*. IEEE. doi:10.1109/cns.2014.6997534.

Samalikova, J. (2012). Process mining application in software process assessment. 10.6100/ir739990

Samalikova, J., Kusters, R., Trienekens, J., Weijters, T., & Siemons, P. (2010). Toward objective software process information: experiences from a case study. *Software Quality Journal, 19*(1), 101–120. doi:10.1007/s11219-010-9105-8.

Samalikova, J., Kusters, R. J., Trienekens, J. J. M., & Weijters, A. J. M. M. (2014). Process mining support for capability maturity model integration-based software process assessment, in principle and in practice. *Journal of Software: Evolution and Process, 26*(7), 714–728. doi:10.1002/smr.1645.

Šamalíková, J., Trienekens, J. J. M., Kusters, R. J., & Weijters, A. J. M. M. (2009). Discovering changes of the change control board process during a software development project using process mining. In *Communications in computer and information science* (pp. 128–136). Springer Berlin Heidelberg. doi:10.1007/978-3-642-04133-4_11.

de San Pedro, J., & Cortadella, J. (2016). Mining structured petri nets for the visualization of process behavior. In *Proceedings of the 31st annual ACM symposium on applied computing - SAC '16*. ACM Press. doi:10.1145/2851613.2851645.

Sánchez-Garzón, I., González-Ferrer, A., & Fernández-Olivares, J. (2013). A knowledge-based architecture for the management of patient-focused care pathways. *Applied Intelligence, 40*(3), 497–524. doi:10.1007/s10489-013-0466-0.
dos Santos, C. F., Piechnicki, F., de Freitas Rocha Loures, E., & Santos, E. A. P. (2017). Mapping the conceptual relationship among data analysis, knowledge generation and decision-making in industrial processes. *Procedia Manufacturing, 11*, 1751–1758. doi:10.1016/j.promfg.2017.07.305.
Santos, R. M. S., Oliveira, T. C., & e Abreu, F. B. (2015). Mining software development process variations. In *Proceedings of the 30th annual acm symposium on applied computing*. In *SAC '15* (pp. 1657–1660). New York, NY, USA: ACM. doi:10.1145/2695664.2696046.
dos Santos Rocha, R., & Fantinato, M. (2013). The use of software product lines for business process management: A systematic literature review. *Information and Software Technology, 55*(8), 1355–1373. doi:10.1016/j.infsof.2013.02.007.
Sarno, R., & Sinaga, F. P. (2015). Business process anomaly detection using ontology-based process modelling and multi-level class association rule learning. *2015 international conference on computer, control, informatics and its applications (IC3ina)*. IEEE. doi:10.1109/ic3ina.2015.7377738.
Schoknecht, A., Thaler, T., Fettke, P., Oberweis, A., & Laue, R. (2017). Similarity of business process models—a state-of-the-art analysis. *ACM Computing Surveys, 50*(4), 1–33. doi:10.1145/3092694.
Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior, 28*(4), 1321–1331. doi:10.1016/j.chb.2012.02.016.
Schwenke, C., Wagner, T., Gellrich, A., & Kabitzsch, K. (2012). Event-based recognition and source identification of transient tailbacks in manufacturing plants. In *Proceedings title: Proceedings of the 2012 winter simulation conference (WSC)*. IEEE. doi:10.1109/wsc.2012.6465013.
Sebu, M. L., & Ciocarlie, H. (2014). Applied process mining in software development. In *2014 ieee 9th ieee international symposium on applied computational intelligence and informatics (saci)* (pp. 55–60). doi:10.1109/SACI.2014.6840098.
Sebu, M. L., & Ciocarlie, H. (2015). Business activity monitoring solution to detect deviations in business process execution. *2015 IEEE 10th jubilee international symposium on applied computational intelligence and informatics*. IEEE. doi:10.1109/saci.2015.7208243.
Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2018). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*. doi:10.1016/j.chb.2018.05.004.
Sedrakyan, G., & Snoeck, M. (2017). Cognitive feedback and behavioral feedforward automation perspectives for modeling and validation in a learning context. In *Communications in computer and information science* (pp. 70–92). Springer International Publishing. doi:10.1007/978-3-319-66302-9_4.
Sedrakyan, G., Snoeck, M., & Weerdt, J. D. (2014). Process mining analysis of conceptual modeling behavior of novices - empirical study using jmermaid modeling and experimental logging environment. *Computers in Human Behavior, 41*, 486–503. doi:10.1016/j.chb.2014.09.054.
Sedrakyan, G., Weerdt, J. D., & Snoeck, M. (2016). Process-mining enabled feedback: "tell me what i did wrong" vs. "tell me how to do it right". *Computers in Human Behavior, 57*, 352–376. doi:10.1016/j.chb.2015.12.040.
Senderovich, A., Francescomarino, C. D., & Maggi, F. M. (2019). From knowledge-driven to data-driven inter-case feature encoding in predictive process monitoring. *Information Systems*. doi:10.1016/j.is.2019.01.007.
Sengupta, B., Banerjee, N., Bisdikian, C., & Hurley, P. (2008). Tracking transaction footprints for non-intrusive end-to-end monitoring. *Cluster Computing, 12*(1), 59–72. doi:10.1007/s10586-008-0066-7.
Shah, V., Khadke, C., & Rana, S. (2015a). Knowledge transition: Discovering workflow models from functional tests. *2015 IEEE international symposium on software reliability engineering workshops (ISSREW)*. IEEE. doi:10.1109/issrew.2015.7392037.
Shah, V., Khadke, C., & Rana, S. (2015b). Mining process models and architectural components from test cases. *2015 IEEE eighth international conference on software testing, verification and validation workshops (ICSTW)*. IEEE. doi:10.1109/icstw.2015.7107428.
Shao, Q., Chen, Y., Tao, S., Yan, X., & Anerousis, N. (2008). Efficient ticket routing by resolution sequence mining. *Proceeding of the 14th ACM SIGKDD international conference on knowledge discovery and data mining - KDD 08*. ACM Press. doi:10.1145/1401890.1401964.
Siddiqui, Z. F., Krempl, G., Spiliopoulou, M., Peña, J. M., Paul, N., & Maestu, F. (2015). Predicting the post-treatment recovery of patients suffering from traumatic brain injury (TBI). *Brain Informatics, 2*(1), 33–44. doi:10.1007/s40708-015-0010-6.
da Silva, M. J., Pereira, C. E., & Götz, M. (2016). Knowledge awareness in alarm system database. *IFAC-PapersOnLine, 49*(28), 297–302. doi:10.1016/j.ifacol.2016.11.051.
Sobocinski, M., Malmberg, J., & Järvelä, S. (2017). Exploring temporal sequences of regulatory phases and associated interactions in low- and high-challenge collaborative learning sessions. *Metacognition and Learning, 12*(2), 275–294. doi:10.1007/s11409-016-9167-5.
Soffer, P., Hinze, A., Koschmider, A., Ziekow, H., Ciccio, C. D., Koldehofe, B., et al. (2019). From event streams to process models and back: Challenges and opportunities. *Information Systems, 81*, 181–200. doi:10.1016/j.is.2017.11.002.
Song, M., & van der Aalst, W. M. (2008). Towards comprehensive support for organizational mining. *Decision Support Systems, 46*(1), 300–317. doi:10.1016/j.dss.2008.07.002.
Song, M., Yang, H., Siadat, S., & Pechenizkiy, M. (2013). A comparative study of dimensionality reduction techniques to enhance trace clustering performances. *Expert Systems with Applications, 40*(9), 3722–3737. doi:10.1016/j.eswa.2012.12.078.
Song, W., Jacobsen, H.-A., Ye, C., & Ma, X. (2016). Process discovery from dependence-complete event logs. *IEEE Transactions on Services Computing, 9*(5), 714–727. doi:10.1109/tsc.2015.2426181.
Southavilay, V., Yacef, K., & Calvo, R. A. (2010). Analysis of collaborative writing processes using hidden markov models and semantic heuristics. In *2010 ieee international conference on data mining workshops* (pp. 543–548). doi:10.1109/ICDMW.2010.118.
van der Spoel, S., van Keulen, M., & Amrit, C. (2013). Process prediction in noisy data sets: A case study in a dutch hospital. In *Lecture notes in business information processing* (pp. 60–83). Springer Berlin Heidelberg. doi:10.1007/978-3-642-40919-6_4.
Stuht, T., Speck, A., Feja, S., Witt, S., & Pulvermüller, E. (2012). Rule determination and process verification using business capabilities. In *Lecture notes in business information processing* (pp. 46–60). Springer Berlin Heidelberg. doi:10.1007/978-3-642-34549-4_4.
Su, Y., & Al-Hakim, L. (2010). Intelligent control model for checking data quality in hospital process management. *2010 the 2nd international conference on industrial mechatronics and automation*. IEEE. doi:10.1109/icindma.2010.5538292.
Subramaniam, S., Kalogeraki, V., Gunopulos, D., Casati, F., Castellanos, M., Dayal, U., et al. (2007). Improving process models by discovering decision points. *Information Systems, 32*(7), 1037–1055. doi:10.1016/j.is.2006.11.001.
Sunindyo, W. D., & Ekaputra, F. J. (2013). OSMF: A framework for OSS process measurement. In *Lecture notes in computer science* (pp. 71–80). Springer Berlin Heidelberg. doi:10.1007/978-3-642-36818-9_8.
Sunindyo, W. D., Moser, T., Winkler, D., & Biffl, S. (2010a). Foundations for event-based process analysis in heterogeneous software engineering environments. *2010 36th EUROMICRO conference on software engineering and advanced applications*. IEEE. doi:10.1109/seaa.2010.52.
Sunindyo, W. D., Moser, T., Winkler, D., & Biffl, S. (2010b). A process model discovery approach for enabling model interoperability in signal engineering. In *Proceedings of the first international workshop on model-drive interoperability - MDI '10*. ACM Press. doi:10.1145/1866272.1866276.
Suriadi, S., Andrews, R., ter Hofstede, A., & Wynn, M. (2017). Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. *Information Systems, 64*, 132–150. doi:10.1016/j.is.2016.07.011.
Sutrisnowati, R. A., Bae, H., Park, J., & Ha, B.-H. (2013). Learning bayesian network from event logs using mutual information test. *2013 IEEE 6th international conference on service-oriented computing and applications*. IEEE. doi:10.1109/soca.2013.38.
Sutrisnowati, R. A., Bae, H., & Song, M. (2015). Bayesian network construction from event log for lateness analysis in port logistics. *Computers & Industrial Engineering, 89*, 53–66. doi:10.1016/j.cie.2014.11.003.
Sutrisnowati, R. A., Yahya, B. N., Bae, H., Pulshashi, I. R., & Adi, T. N. (2017). Scalable indexing algorithm for multi-dimensional time-gap analysis with distributed computing. *Procedia Computer Science, 124*, 224–231. doi:10.1016/j.procs.2017.12.150.
Tesanovic, A., Manev, G., Pechenizkiy, M., & Vasilyeva, E. (2009). eHealth personalization in the next generation RPM systems. *2009 22nd IEEE international symposium on computer-based medical systems*. IEEE. doi:10.1109/cbms.2009.5255383.
Toth, K., Machalik, K., Fogarassy, G., & Vathy-Fogarassy, A. (2017). Applicability of process mining in the exploration of healthcare sequences. *2017 IEEE 30th Neumann colloquium (NC)*. IEEE. doi:10.1109/nc.2017.8263273.
Toyawanit, T., & Premchaiswadi, W. (2016). Applying inductive visual miner technique to analyze and detect problems in procedures of a hospital in thailand. *2016 14th international conference on ICT and knowledge engineering (ICT&KE)*. IEEE. doi:10.1109/ictke.2016.7804105.
Trcka, N., & Pechenizkiy, M. (2009). From local patterns to global models: Towards domain driven educational process mining. In *2009 ninth international conference on intelligent systems design and applications* (pp. 1114–1119). doi:10.1109/ISDA.2009.159.
Trcka, N., Pechenizkiy, M., & van der Aalst, W. M. P. (2011). Chapter 9: Process mining from educational data. *Edm 2011*.
Tsumoto, S., & Hirano, S. (2015). Analytics for hospital management. In *Proceedings of the ase bigdata & socialinformatics 2015*. In *ASE BD&SI '15* (pp. 15:1–15:6). New York, NY, USA: ACM. doi:10.1145/2818869.2818894.
Tsumoto, S., & Hirano, S. (2016). Data mining oriented software quality estimation. *Procedia Computer Science, 91*, 1028–1037. doi:10.1016/j.procs.2016.07.141.
Tu, T. B. H., & Song, M. (2016). Analysis and prediction cost of manufacturing process based on process mining. *2016 international conference on industrial engineering, management science and application (ICIMSA)*. IEEE. doi:10.1109/icimsa.2016.7503993.
Tupa, J. (2008). Electronics production and process controlling. *2008 2nd electronics systemintegration technology conference*. IEEE. doi:10.1109/estc.2008.4684441.
Valle, A. M., Santos, E. A., & Loures, E. R. (2017). Applying process mining techniques in software process appraisals. *Information and Software Technology, 87*, 19–31. doi:10.1016/j.infsof.2017.01.004.
Vanderfeesten, I., Reijers, H. A., & van der Aalst, W. M. (2011). Product-based workflow support. *Information Systems, 36*(2), 517–535. doi:10.1016/j.is.2010.09.008.
Vasilyev, E., Ferreira, D. R., & Iijima, J. (2013). Using inductive reasoning to find the cause of process delays. *2013 IEEE 15th conference on business informatics*. IEEE. doi:10.1109/cbi.2013.41.

Vasyutynskyy, V., Gellrich, A., Kabitzsch, K., & Wustmann, D. (2010). Analysis of internal logistic systems based on event logs. *2010 IEEE 15th conference on emerging technologies & factory automation (ETFA 2010)*. IEEE. doi:10.1109/etfa.2010.5641202.

Viale, P., Frydman, C., & Pinaton, J. (2011). New methodology for modeling large scale manufacturing process: Using process mining methods and experts' knowledge. *2011 9th IEEE/ACS international conference on computer systems and applications (AICCSA)*. IEEE. doi:10.1109/aiccsa.2011.6126595.

Viale, P., Giambiasi, N., Frydman, C., & Pinaton, J. (2012). Validating manufacturing processes using timed sequential machines: Application to a semiconductor manufacturing process (wip). In *Proceedings of the 2012 symposium on theory of modeling and simulation - devs integrative m&s symposium*. In *TMS/DEVS '12* (pp. 51:1–51:6). San Diego, CA, USA: Society for Computer Simulation International.

Vidal, J. C., Vázquez-Barreiros, B., Lama, M., & Mucientes, M. (2016). Recompiling learning processes from event logs. *Knowledge-Based Systems, 100*, 160–174. doi:10.1016/j.knosys.2016.03.003.

Vitali, M., & Pernici, B. (2016). Interconnecting processes through IoT in a healthcare scenario. *2016 IEEE international smart cities conference (ISC2)*. IEEE. doi:10.1109/isc2.2016.7580760.

Vogelgesang, T., & Appelrath, H.-J. (2013). Multidimensional process mining. In *Proceedings of the joint EDBT/ICDT 2013 workshops on - EDBT '13*. ACM Press. doi:10.1145/2457317.2457321.

Wagner, T., Schwenke, C., & Kabitzsch, K. (2012). Modeling and wafer defect analysis in semiconductor automated material handling systems. In *Proceedings of the 2012 winter simulation conference (wsc)* (pp. 1–12). doi:10.1109/WSC.2012.6465006.

qiong Wang, H., shu Zhou, T., li Tian, L., ming Qian, Y., & song Li, J. (2014). Creating hospital-specific customized clinical pathways by applying semantic reasoning to clinical data. *Journal of Biomedical Informatics, 52*, 354–363. doi:10.1016/j.jbi.2014.07.017.

Wang, J., Jiang, L., & Cai, H. (2014). Scenario-based method for business process analysis and improvement in SOA. *2014 IEEE 11th international conference on e-business engineering*. IEEE. doi:10.1109/icebe.2014.16.

Wang, J., Tan, S., Wen, L., Wong, R. K., & Guo, Q. (2012). An empirical evaluation of process mining algorithms based on structural and behavioral similarities. In *Proceedings of the 27th annual ACM symposium on applied computing - SAC '12*. ACM Press. doi:10.1145/2245276.2245316.

Wang, J., Wong, R. K., Ding, J., Guo, Q., & Wen, L. (2012). On recommendation of process mining algorithms. *2012 IEEE 19th international conference on web services*. IEEE. doi:10.1109/icws.2012.52.

Wang, J., Wong, R. K., Ding, J., Guo, Q., & Wen, L. (2013). Efficient selection of process mining algorithms. *IEEE Transactions on Services Computing, 6*(4), 484–496. doi:10.1109/tsc.2012.20.

Wang, J., Zhu, B., Wang, Y., & Huang, L. (2016). Mining organizational behaviors in collaborative logistics chain: An empirical study in a port. *2016 international conference on logistics, informatics and service sciences (LISS)*. IEEE. doi:10.1109/liss.2016.7854453.

Wang, T., Tian, X., Yu, M., Qi, X., & Yang, L. (2017). Stage division and pattern discovery of complex patient care processes. *Journal of Systems Science and Complexity, 30*(5), 1136–1159. doi:10.1007/s11424-017-5302-x.

Wang, Y., Caron, F., Vanthienen, J., Huang, L., & Guo, Y. (2014). Acquiring logistics process intelligence: Methodology and an application for a chinese bulk port. *Expert Systems with Applications, 41*(1), 195–209. doi:10.1016/j.eswa.2013.07.021.

Wang, Y., White, W. M., & Andersen, E. (2017). PathViewer: Visualizing pathways through student data. In *Proceedings of the 2017 CHI conference on human factors in computing systems - CHI '17*. ACM Press. doi:10.1145/3025453.3025819.

Wang, Z., ter Hofstede, A. H., Ouyang, C., Wynn, M., Wang, J., & Zhu, X. (2014). How to guarantee compliance between workflows and product lifecycles? *Information Systems, 42*, 195–215. doi:10.1016/j.is.2014.01.006.

xia Wang, Z., min Wang, J., chen Zhu, X., & jie Wen, L. (2012). Verification of workflow nets with transition conditions. *Journal of Zhejiang University SCIENCE C, 13*(7), 483–509. doi:10.1631/jzus.c1100364.

Wang, Z., Yao, Q., & Sun, Y. (2012). The research of process mining assessment used in business intelligence. *2012 IEEE/ACIS 11th international conference on computer and information science*. IEEE. doi:10.1109/icis.2012.109.

Weber, I., Nepal, S., & Zhu, L. (2016). Developing dependable and secure cloud applications. *IEEE Internet Computing, 20*(3), 74–79. doi:10.1109/mic.2016.67.

Weerapong, S., Porouhan, P., & Premchaiswadi, W. (2012). Process mining using $\alpha$-algorithm as a tool (a case study of student registration). *2012 tenth international conference on ICT and knowledge engineering*. IEEE. doi:10.1109/ictke.2012.6408558.

Weerdt, J. D., Backer, M. D., Vanthienen, J., & Baesens, B. (2012). A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs. *Information Systems, 37*(7), 654–676. doi:10.1016/j.is.2012.02.004.

Weerdt, J. D., vanden Broucke, S., Vanthienen, J., & Baesens, B. (2013). Active trace clustering for improved process discovery. *IEEE Transactions on Knowledge and Data Engineering, 25*(12), 2708–2720. doi:10.1109/tkde.2013.64.

Weerdt, J. D., Schupp, A., Vanderloock, A., & Baesens, B. (2013). Process mining for the multi-faceted analysis of business processes—a case study in a financial services organization. *Computers in Industry, 64*(1), 57–67. doi:10.1016/j.compind.2012.09.010.

Weidlich, M., Mendling, J., & Weske, M. (2012). Propagating changes between aligned process models. *Journal of Systems and Software, 85*(8), 1885–1898. doi:10.1016/j.jss.2012.02.044.

Weijters, A. J. M. M., & van der Aalst, W. M. P. (2003). Rediscovering workflow models from event-based data using little thumb. *Integrated Computer-Aided Engineering, 10*(2), 151–162.

Werner, M. (2016). Process model representation layers for financial audits. *2016 49th Hawaii international conference on system sciences (HICSS)*. IEEE. doi:10.1109/hicss.2016.659.

Werner, M. (2017). Financial process mining - Accounting data structure dependent control flow inference. *International Journal of Accounting Information Systems, 25*, 57–80. doi:10.1016/j.accinf.2017.03.004.

Werner, M., & Gehrke, N. (2015). Multilevel process mining for financial audits. *IEEE Transactions on Services Computing, 8*(6), 820–832. doi:10.1109/tsc.2015.2457907.

Werner, M., Gehrke, N., & Nuttgens, M. (2012). Business process mining and reconstruction for financial audits. *2012 45th Hawaii international conference on system sciences*. IEEE. doi:10.1109/hicss.2012.141.

Wil, M. P., van der Aalst, B. V. D. E. K., Rubin, V. A., & Günther, C. (2006). *Process mining: A two-step approach using transition systems and regions*.

Wolf, H., Herrmann, K., & Rothermel, K. (2013). Dealing with uncertainty. *ACM Transactions on Intelligent Systems and Technology, 4*(4), 1–23. doi:10.1145/2508037.2508046.

Worgan, S. F., Behera, A., Cohn, A. G., & Hogg, D. C. (2011). Exploiting petri-net structure for activity classification and user instruction within an industrial setting. In *Proceedings of the 13th international conference on multimodal interfaces - ICMI '11*. ACM Press. doi:10.1145/2070481.2070502.

Wu, S., Lai, C., & Sun, Y. (2006). Applying data mining and petri net in reengineering of manufacture management information system. *2006 IEEE asia-pacific conference on services computing (APSCC'06)*. IEEE. doi:10.1109/apscc.2006.36.

Xing, J., Li, Z., Cheng, Y., Yin, F., Li, B., & Chen, L. (2007). Mining process models from event logs in distributed bioinformatics workflows. *The first international symposium on data, privacy, and e-commerce (ISDPE 2007)*. IEEE. doi:10.1109/isdpe.2007.129.

Xu, D., Xu, W., Bavikati, B. K., & Wong, W. E. (2012). Mining executable specifications of web applications from selenium IDE tests. *2012 IEEE sixth international conference on software security and reliability*. IEEE. doi:10.1109/sere.2012.39.

Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: State of the art and future trends. *International Journal of Production Research, 56*(8), 2941–2962. doi:10.1080/00207543.2018.1444806.

Xu, X., Jin, T., & Wang, J. (2016). Summarizing patient daily activities for clinical pathway mining. *2016 IEEE 18th international conference on e-health networking, applications and services (healthcom)*. IEEE. doi:10.1109/healthcom.2016.7749453.

Xu, X., Jin, T., Wei, Z., Lv, C., & Wang, J. (2016). TCPM: Topic-based clinical pathway mining. *2016 IEEE first international conference on connected health: Applications, systems and engineering technologies (CHASE)*. IEEE. doi:10.1109/chase.2016.17.

Xu, X., Jin, T., Wei, Z., & Wang, J. (2017). Incorporating domain knowledge into clinical goal discovering for clinical pathway mining. *2017 IEEE EMBS international conference on biomedical & health informatics (BHI)*. IEEE. doi:10.1109/bhi.2017.7897255.

Xu, X., Weber, I., Bass, L., Zhu, L., Wada, H., & Teng, F. (2013). Detecting cloud provisioning errors using an annotated process model. In *Proceedings of the 8th workshop on middleware for next generation internet computing - MW4nextgen '13*. ACM Press. doi:10.1145/2541608.2541613.

Xu, X., Zhu, L., Weber, I., Bass, L., & Sun, D. (2014). POD-diagnosis: Error diagnosis of sporadic operations on cloud applications. *2014 44th annual IEEE/IFIP international conference on dependable systems and networks*. IEEE. doi:10.1109/dsn.2014.94.

Yahya, B. N., Song, M., Bae, H., ook Sul, S., & Wu, J.-Z. (2016). Domain-driven actionable process model discovery. *Computers & Industrial Engineering, 99*, 382–400. doi:10.1016/j.cie.2016.05.010.

Yampaka, T., & Chongstitvatana, P. (2016). An application of process mining for queueing system in health service. *2016 13th international joint conference on computer science and software engineering (JCSSE)*. IEEE. doi:10.1109/jcsse.2016.7748865.

Yang, H., Park, M., Cho, M., Song, M., & Kim, S. (2014). A system architecture for manufacturing process analysis based on big data and process mining techniques. *2014 IEEE international conference on big data (big data)*. IEEE. doi:10.1109/bigdata.2014.7004336.

Yang, S., Dong, X., Sun, L., Zhou, Y., Farneth, R. A., Xiong, H., et al. (2017). A data-driven process recommender framework. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining - KDD '17*. ACM Press. doi:10.1145/3097983.3098174.

Yang, S., Zhou, M., Chen, S., Dong, X., Ahmed, O., Burd, R. S., et al. (2017). Medical workflow modeling using alignment-guided state-splitting HMM. *2017 IEEE international conference on healthcare informatics (ICHI)*. IEEE. doi:10.1109/ichi.2017.66.

Yang, W., & Su, Q. (2014). Process mining for clinical pathway: Literature review and future directions. *2014 11th international conference on service systems and service management (ICSSSM)*. IEEE. doi:10.1109/icsssm.2014.6943412.

Yano, K., Nomura, Y., & Kanai, T. (2013). A practical approach to automated business process discovery. *2013 17th IEEE international enterprise distributed object computing conference workshops*. IEEE. doi:10.1109/edocw.2013.13.

Yao, J., Shepherd, M., Zhou, J., Fu, L., Li, F., Quebe, D., et al. (2015). Guided analytic workflows through service composition for population health studies. *2015 IEEE international conference on services computing*. IEEE. doi:10.1109/scc.2015.99.

Yi, J., & Lai, C. (2008). Research on reengineering of ERP system based on data mining and MAS. *2008 international symposium on knowledge acquisition and modeling*. IEEE. doi:10.1109/kam.2008.117.

Yoo, S., Cho, M., Kim, E., Kim, S., Sim, Y., Yoo, D., et al. (2016). Assessment of hospital processes using a process mining technique: Outpatient process analysis at a tertiary hospital. *International Journal of Medical Informatics, 88*, 34–43. doi:10.1016/j.ijmedinf.2015.12.018.

Yoo, S., Jung, S. Y., Kim, S., Kim, E., Lee, K.-H., Chung, E., et al. (2016). A personalized mobile patient guide system for a patient-centered smart hospital: Lessons learned from a usability test and satisfaction survey in a tertiary university hospital. *International Journal of Medical Informatics, 91*, 20–30. doi:10.1016/j.ijmedinf.2016.04.003.

Yu, D., & Wang, J. (2017). Mining collaboration patterns of software development processes based on trace alignment. In *Proceedings of the 2017 international conference on software and system process - ICSSP 2017*. ACM Press. doi:10.1145/3084100.3084103.

van Zelst, S. J., van Dongen, B. F., & van der Aalst, W. M. P. (2017). Event stream-based process discovery using abstract representations. *Knowledge and Information Systems, 54*(2), 407–435. doi:10.1007/s10115-017-1060-2.

Zeng, Q., Lu, F., Liu, C., Duan, H., & Zhou, C. (2015). Modeling and verification for cross-department collaborative business processes using extended petri nets. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45*(2), 349–362. doi:10.1109/tsmc.2014.2334276.

Zeng, Q., Sun, S. X., Duan, H., Liu, C., & Wang, H. (2013). Cross-organizational collaborative workflow mining from a multi-source log. *Decision Support Systems, 54*(3), 1280–1301. doi:10.1016/j.dss.2012.12.001.

Zerouali, A., & Mens, T. (2017). An empirical comparison of the development history of cloudstack and eucalyptus. In *Proceedings of the 2017 international conference on smart digital environment - ICSDE '17*. ACM Press. doi:10.1145/3128128.3128146.

Zha, H., van der Aalst, W. M. P., Wang, J., Wen, L., & Sun, J. (2011). Verifying workflow processes: A transformation-based approach. *Software & Systems Modeling, 10*(2), 253–264. doi:10.1007/s10270-010-0149-9.

Zha, H., Wang, J., Wen, L., & Wang, C. (2009). A label-free similarity measure between workflow nets. *2009 IEEE asia-pacific services computing conference (APSCC)*. IEEE. doi:10.1109/apscc.2009.5394086.

Zhang, M., & Hochstein, L. (2009). Fitting a workflow model to captured development data. *2009 3rd international symposium on empirical software engineering and measurement*. IEEE. doi:10.1109/esem.2009.5315990.

Zhang, X., & Chen, S. (2012). Pathway identification via process mining for patients with multiple conditions. *2012 IEEE international conference on industrial engineering and engineering management*. IEEE. doi:10.1109/ieem.2012.6838048.

Zhang, Y., Martikainen, O., Pulli, P., & Naumov, V. (2011). Real-time process data acquisition with bluetooth. In *Proceedings of the 4th international symposium on applied sciences in biomedical and communication technologies - ISABEL '11*. ACM Press. doi:10.1145/2093698.2093719.

Zhang, Y., Padman, R., & Patel, N. (2015). Paving the COWpath: Learning and visualizing clinical pathways from electronic health record data. *Journal of Biomedical Informatics, 58*, 186–197. doi:10.1016/j.jbi.2015.09.009.

Zhao, W., Liu, H., Dai, W., & Ma, J. (2015). An entropy-based clustering ensemble method to support resource allocation in business process management. *Knowledge and Information Systems, 48*(2), 305–330. doi:10.1007/s10115-015-0879-7.

Zhao, W., Zeng, Q., Zheng, G., & Yang, L. (2017). The resource allocation model for multi-process instances based on particle swarm optimization. *Information Systems Frontiers, 19*(5), 1057–1066. doi:10.1007/s10796-017-9743-5.

Zhaoxia, W., Jianmin, W., Lijie, W., & Yingbo, L. (2009). Extracting real-life log with data perspective in PLM system for business process analysis. *2009 WRI world congress on computer science and information engineering*. IEEE. doi:10.1109/csie.2009.645.

Zhou, M., Yang, S., Li, X., Lv, S., Chen, S., Marsic, I., et al. (2017). Evaluation of trace alignment quality and its application in medical process mining. *2017 IEEE international conference on healthcare informatics (ICHI)*. IEEE. doi:10.1109/ichi.2017.57.

Zhou, W., & Piramuthu, S. (2010). Framework, strategy and evaluation of health care processes with RFID. *Decision Support Systems, 50*(1), 222–233. doi:10.1016/j.dss.2010.08.003.

Zhou, Z., Wang, Y., & Li, L. (2014). Process mining based modeling and analysis of workflows in clinical care - A case study in a chicago outpatient clinic. In *Proceedings of the 11th IEEE international conference on networking, sensing and control*. IEEE. doi:10.1109/icnsc.2014.6819692.