Process Mining: A Guide for Practitioners

Fredrik Milani[1(✉)], Katsiaryna Lashkevich[1], Fabrizio Maria Maggi[2],
and Chiara Di Francescomarino[3]

[1] University of Tartu, Tartu, Estonia
{milani,katsiaryna.lashkevich}@ut.ee
[2] Free University of Bozen-Bolzano, Bolzano, Italy
maggi@inf.unibz.it
[3] FBK-IRST, Trento, Italy
dfmchiara@fbk.eu

Abstract. In the last years, process mining has significantly matured and has increasingly been applied by companies in industrial contexts. However, with the growing number of process mining methods, practitioners might find it difficult to identify which ones to apply in specific contexts and to understand the specific business value of each process mining technique. This paper's main objective is to develop a business-oriented framework capturing the main process mining use cases and the business-oriented questions they can answer. We conducted a Systematic Literature Review (SLR) and we used the review and the extracted data to develop a framework that (1) classifies existing process mining use cases connecting them to specific methods implementing them, and (2) identifies business-oriented questions that process mining use cases can answer. Practitioners can use the framework to navigate through the available process mining use cases and to identify the process mining methods suitable for their needs.

Keywords: Business process management · Process mining · Use cases · Systematic Literature Review · Business value

1 Introduction

A business process is a set of activities executed in a given setting to achieve predefined business objectives [25]. Since business processes constitute the operational foundation of organizations, companies seek to manage and improve them. Business processes are commonly supported by information systems that record data on the executions of the processes. These records are referred to as event logs. An event log consists of traces that capture the execution of a business process instance (a.k.a. a case). A case consists of a sequence of time-stamped events, each representing the execution of an activity. Process mining is a family of methods that analyze business processes based on their observed behavior recorded in event logs.

In the past two decades, research on process mining has made advances resulting in the generation of a large corpus of academic literature [63]. However,

the number of studies on process mining methods might be disconcerting when companies seek to understand how process mining can be applied for improving business processes. More specifically, companies might find it challenging to understand (1) which prominent process mining use cases are available and (2) what business-oriented questions such use cases answer.

As such, the objective of this study is to develop, based on a Systematic Literature Review (SLR), a business-oriented framework that classifies existing process mining use cases relating them with specific methods and with the business-oriented questions they can answer. Therefore, the framework can aid practitioners that seek to use data-driven approaches to manage their business processes in exploring how process mining methods can add value to their business and what process-related analysis such methods can support. Thus, the main research questions we seek to examine in this paper are *What are the main use cases for existing process mining methods?* and *What business-oriented questions do existing process mining use cases answer?*. We conducted the SLR following the guidelines proposed in [45], retrieved 2293 papers using a keyword-based search from electronic libraries, and filtered them according to predefined inclusion and exclusion criteria. We finally identified a corpus of 839 relevant papers that we reviewed. Then, we used the review and the extracted data to develop a business-oriented a framework that could represent a valid instrument to guide companies in how process mining can support their business.

The remainder of the paper is organized as follows. In Sect. 2, we position our work against related work. Section 3 presents the SLR protocol. Section 4 summarizes the results, and Sect. 5 presents and discusses the framework. Finally, we conclude the paper in Sect. 6.

2   Related Work

In this section, we position our work against existing reviews within the process mining field. More specifically, we consider systematic mapping studies, process mining reviews in specific industrial sectors, and reviews on how process mining is applied in industry.

In [63], dos Santos Garcia et al. present a systematic mapping study providing an overview of the main process mining branches, algorithms, and application domains. A similar study is discussed by Maita et al. in [54]. These studies highlight that most of the process mining publications can be associated with process model discovery. This is in line with our findings. However, while these mapping studies focus on the current state of the process mining research, we examined empirically validated process mining methods to elicit how they might deliver value to companies. Thus, we take a business-oriented perspective allowing practitioners to link the everyday issues they have to deal with in their organizations with the process mining techniques that can help them solving these issues. Furthermore, Maita et al. classified process mining techniques into 3 main branches, i.e., process model discovery, conformance checking, and process model enhancement as proposed in [1]. This classification is also applied

in other studies, such as [18]. However, the application of process mining has evolved in recent years. Therefore, our framework extends this classification by incorporating more recent process mining use cases.

Several reviews have also been conducted within specific industry sectors. For instance, literature reviews have been conducted with focus on healthcare [6,29,33,36,61], educational processes [34], and supply chain [41]. While these reviews focus on a specific application domain, we consider business-oriented questions that are answered at a domain-agnostic level.

Finally, reviews have also been conducted on how organizations use process mining. For instance, Thiede et al. [70] show with their study that process mining is not sufficiently leveraged in the context of cross-system and cross-organizational processes. Corallo et al. [16] primarily provide an overview of software tools that support process mining analysis in industry. Eggers and Hein [27], instead, examine capabilities and practices required to enable the realization of value using process mining in an organization. These studies provide valuable insights on different aspects of how process mining is applied in industry, but they do not guide practitioners in selecting use cases and methods to answer common business-oriented questions.

3   Systematic Literature Review

Our research objective is to develop a framework for classifying process mining use cases and the business issues they might address. This objective is achieved by answering two research questions (RQ). The first research question (RQ1) aims at identifying and categorizing process mining use cases, such as conformance checking and predictive monitoring. Therefore, RQ1 is formulated as *What are the main use cases for existing process mining methods?* The second research question (RQ2) aims at eliciting the business-oriented questions that the outputs of process mining methods can answer. Therefore, RQ2 is formulated as *What business-oriented questions do existing process mining use cases answer?* To answer these questions, we employ the SLR method as it allows us to collect relevant studies and, based on the review of existing research, to develop a framework for classifying them [15]. We followed the guidelines proposed by Kitchenham [45] according to which an SLR has three consecutive phases: planning, execution, and reporting.

Our SLR is composed of two parts. The first one (SLR review) aims at identifying other SLR studies on specific process mining use cases (e.g., [5] for process model discovery and [26] for conformance checking). The list of final papers in each of these SLR studies was extracted. However, as not all process mining use cases have a dedicated SLR study, we conducted a second review (PM review) targeting papers that have applied process mining techniques to real-life event logs. The lists of final papers retrieved from each SLR study identified with the SLR review was combined with the hits obtained with the PM review. The merged list of candidate papers was subjected to content screening. We followed the same guidelines for both parts, i.e., we developed search strings, identified

search sources, filtered the results according to predefined criteria, identified additional relevant papers through backwards referencing, and extracted data according to a predefined form. Below, we provide a summary of these steps. The review protocol[1] and the list of final papers[2] are available online.

The planning phase of our SLR includes developing search strings, identifying search sources, defining selection criteria, and defining the data extraction strategy. We derived the search strings from our research questions as suggested in [45]. For the SLR review, the aim was to capture SLR studies on process mining. Therefore, we used the search string *"process mining" AND "systematic literature review"*. The search strings for the PM review were, instead, derived from the research questions and included the terms "process mining", "workflow mining" (as this term is sometimes used interchangeably with "process mining"), "real-life", "real-world", and "case study". Then, we applied the search strings to electronic databases. We selected Scopus and Web of Science for both parts as they index the venues where most research on process mining is published.

We then defined the selection criteria to identify relevant studies. These criteria, expressed as exclusion (EC) and inclusion (IC) criteria, allowed us to filter the initial list of papers to keep only those that are relevant to answer the research questions. For the SLR review, duplicate papers (EC1), papers not written in English (EC2), and papers inaccessible via the digital libraries subscribed by the University of Tartu, or otherwise unavailable (EC3) were excluded. In addition, with the first inclusion criterion (IC1), we filtered out studies that were not specifically about process mining and the second one (IC2) served to identify studies that applied SLR to identify relevant papers for specific process mining use cases. Thus, studies focusing on, for instance, evaluating process mining algorithms, such as [57] were excluded.

We applied the three exclusion criteria above also for the PM review. In addition, papers having less than five pages were discarded (EC4). The list of candidate papers was then filtered based on the inclusion criteria. We first excluded papers not within the domain of process mining (IC1). Then, we identified papers that apply process mining to real-life event logs (IC2). This criterion was included for two reasons. The first one is to identify methods that are applicable to real-life, often challenging, business settings. The second one is to identify papers that address business process aspects existing in real business contexts. The third inclusion criterion (IC3) was aimed at excluding papers that consider process mining for applications unrelated to business, such as [59], where discovery algorithms for managing noisy event logs are compared. Finally, the fourth inclusion criterion filters out papers that do not provide sufficient information to elicit business-oriented questions (IC4).

The final step of planning an SLR is data extraction. The objective of this step is to define the data extraction form to reduce the opportunity for bias. We developed the data extraction form according to the suggestions provided in [31,60]. The data extraction form consists of two parts. The first part was used to

---

[1] https://doi.org/10.6084/m9.figshare.17099462.
[2] https://doi.org/10.6084/m9.figshare.12933239.

extract the metadata of the papers, i.e., paper id, title, authors, and publication year. In the second part, the data extracted concerned process mining use cases (such as process model discovery or performance analysis) and the questions being answered by the process mining methods.

We conducted both searches in February 2020. A summary of the detailed procedure applied is depicted in Fig. 1. For the SLR review, the application of the search string to the electronic databases returned 132 hits from Scopus and 61 from Web of Science, making it a total of 193 candidate papers. After having applied the exclusion criteria, 60 papers remained. Of these, 15 were removed as they did not meet IC1, resulting in 45 papers. The application of IC2 filtered out additional 26 papers, resulting in 19 papers left. Finally, backward and forward referencing identified two additional papers, resulting in a final list of 21 relevant studies. The data extraction for this part consisted in exporting all studies included in the final lists of all 21 relevant SLR publications. These were merged with the hits resulting from the search conducted for the PM review. From the 21 SLR studies, a total of 702 papers were extracted.
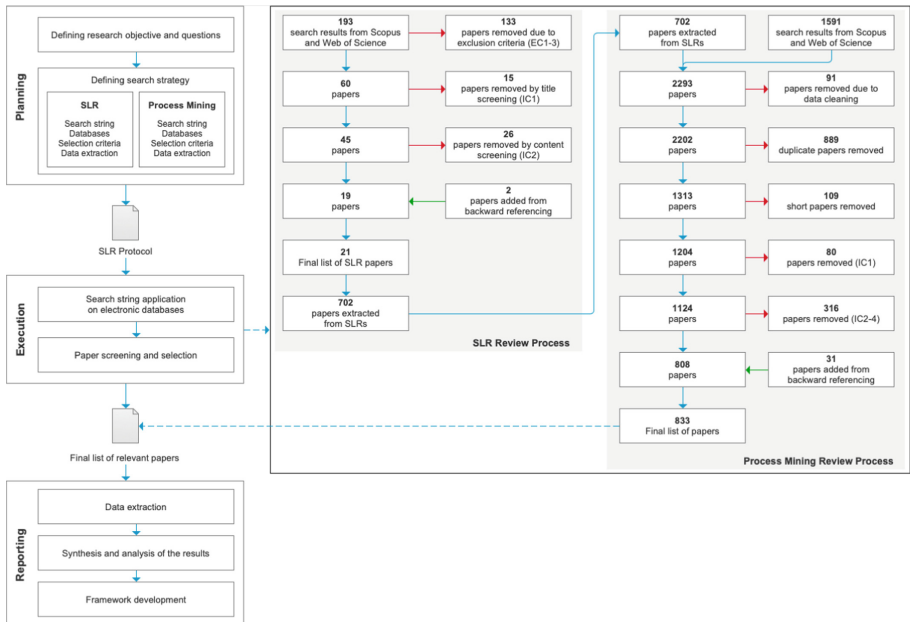

Fig. 1. Application of the SLR protocol.

For the PM review, applying the search strings resulted in 1021 hits from Scopus and 570 from Web of Science, making it 1591 hits in total. With the 702 papers added from the first part, the total number was 2293. We discarded 91 papers that were unavailable, 889 duplicates and 109 short papers. A total of 1204 papers remained. The first inclusion criterion (IC1) resulted in the removal

of 80 papers. In the second filtering, where IC 2-4 were considered, 316 papers were removed, resulting in 808 papers left. Additional 31 papers were identified from the backward and forward referencing. Thus, the final list consists of 839 relevant papers.

4   Results

In this section, we first present the identified use cases of empirically validated process mining methods. Then, we relate them to the business-oriented questions they answer.
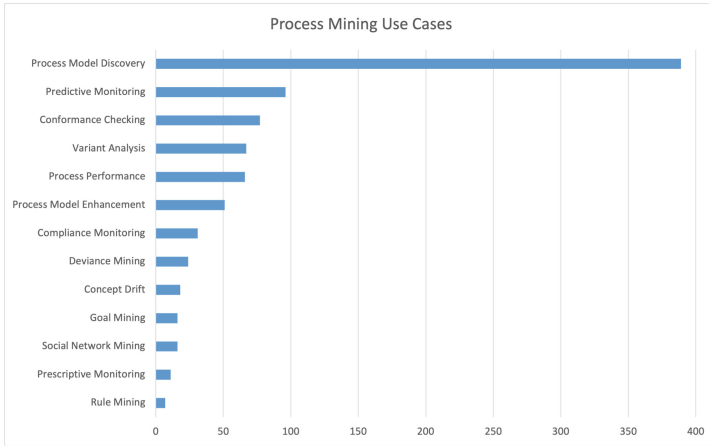
4.1   Process Mining Use Cases

The most common use case identified in the process mining literature is *process model discovery* (see Fig. 2). Business processes are commonly supported by information systems that log information on the process executions. When such logs are available, they can be used to discover process models automatically. Process model discovery takes such event logs as input and produces a process model. Process model discovery is, therefore, used to build procedural process models (e.g., using BPMN or Petri nets) [43], declarative process models (e.g., using the Declare language) [22,51], or hybrid [53] process models containing both a procedural and a declarative part. Use cases related to social network, goal, and rule mining focus on discovering other aspects of the process executions. *Social network mining* analyzes processes from an organizational perspective, i.e., discovers the performers involved in a case and their relations [2]. *Goal mining*, on the other hand, focuses on the process goals [74]. While process model discovery is activity-oriented, goal modeling seeks to discover the process actors' intentions related to the execution of these activities [20]. Finally, *rule mining* [62], also referred to as decision mining, examines the data attributes in an event log to elicit the rules behind the choices made in the process.

Event logs hold information on the actual process executions, which is not necessarily aligned with process models. Therefore, models can be enhanced using *process model enhancement* techniques. In particular, process models can be repaired to better represent the process executions [30,50] or extended with additional data recorded in the event logs [64].

Concept drift identifies changes in the process behavior. The discovery of process models from event logs implicitly assumes that the process model remains stable throughout the time period recorded in the event log. However, this is not always true, as the process might change during the recorded period. Thus, the concept drift use case focuses on detecting changes in the process behavior over time [49].

The second most common use case is *predictive monitoring*. Such use case aims at predicting the outcome of active cases, i.e., cases that are uncompleted and, therefore, still ongoing [24]. Learning from an event log of historical cases, predictive monitoring techniques are able to predict the remaining time of an

Fig. 2. Process mining use cases.

ongoing case [73], delays [65], next activities [66], waiting times [7], outcomes [68], risks [14], costs [72], or performance indicators [17]. Since hyperparameter configuration in predictive process monitoring is crucial and often difficult for users, some works provide methods to support hyperparameter optimization in predictive process monitoring [23]. *Prescriptive monitoring* can be viewed as an extension of predictive monitoring. While predictive monitoring forecasts the likelihood of a case ending up with a desirable/undesirable outcome, it does not suggest the interventions that can increase/reduce the probability of such an outcome. Prescriptive monitoring, on the other hand, seeks to identify specific interventions, such as next activities to be executed [71], resource allocation [75], resource selection [44] to improve the likelihood of a favourable outcome, or when an intervention is needed based on the trade-off cost-gain [69].

The third most common use case we identified is *conformance checking*. Conformance checking aims at examining if the behavior of a process execution, as derived from an event log, conforms with the expected behavior (represented as a process model) [13]. This can be done by simply showing to the user where the process execution deviates from the process model [12,42], or by providing a way to align the deviant process execution with the closest compliant case [3,19,46]. *Compliance monitoring* follows the same principle as conformance checking. However, while conformance checking is applied to completed process cases, compliance monitoring checks whether the behavior of active cases is compliant with predefined rules and constraints [48].

The fourth most common use case we found is *variant analysis*. Executions of a business process commonly include variants, i.e., cases that follow the same path (characterized by the same sequence of activities) [67]. Variant analysis enable identifying these variants in an event log. Variant analysis can also be applied for identifying differences and similarities between different variants [10,38].

Deviance mining, instead, aims at explaining why a certain variant deviates from the most frequently taken path [8, 55].

A last use case concerns assessing *process performance* [56]. The performance measured can be the duration of a process execution [40], the resource utilization [39], or the quality of the products/services provided [4]. The performance of several connected processes can also be assessed [28] and the process performance trends over time [21].

4.2   Business-Oriented Questions

Some process mining methods answer questions that are descriptive. For instance, process model discovery answers the question *How are the cases of a procedural, a declarative, or a hybrid process executed?* The answer is expressed as a process model representing the process behavior as recorded in the event log. Process cases can commonly grouped into variants. Therefore, some process mining methods answer the question *What are the main variants of a process?* Other process mining methods answer, instead, questions to quantitatively describe processes. For instance, process mining can answer questions such as *What is the duration-, resource-, quality-related performance of a case?*

There are also methods answering comparative questions to compare two or more process cases. For instance, variant analysis methods might be used to compare different variants of a process thus answering the questions *What are the similarities between two or more variants of a process?* and *What are the differences between two or more variants of a process?* Similarly, conformance checking seeks to compare the prescribed behavior of a process with the observed behavior, i.e., comparing what a process model stipulates and how the process is executed in reality. Therefore, conformance checking answers questions such as *Where does a case differ from a process model?* Another type of comparative questions are the ones related to how the process behavior changes over time such as *How has the process behavior, or its performance changed over time?*

Process mining also answers questions that seek explanatory answers, i.e., providing information that explains relations among different entities. For instance, some process mining methods answer questions such as *How is the performance of a case affected by other factors?* Likewise, deviance mining provides explanatory information by answering the question *Why do some cases deviate from the normal flow?* Methods that compare a model with historical (conformance checking) or live (compliance monitoring) cases could provide explanatory information by answering questions such as *Given a non-compliant case, what is closest compliant case?* Predictive monitoring methods answer (forecasting) descriptive questions such as *What are the predicted remaining times, delays, next activities, waiting times, outcomes, risks, costs, or performance indicators of an ongoing case?*

Finally, there are process mining methods aiming at providing suggestions on how to redesign a process model to improve understandability, or optimize the likelihood of a favourable outcome of ongoing process executions. For instance,

model repair techniques provide input for improving a process model by answering the recommendatory questions *How can a process model be repaired to better reflect the actual execution of the process?* and *How can the understandability of the mined process models be improved?* On the other hand, prescriptive monitoring methods provide recommendations on how an ongoing process case should be executed to reach a positive outcome. For instance, recommendations can be given on which variant to follow thus answering the question *What is the recommended execution path of an ongoing case?* Recommendations also extend to resources and their allocation by answering questions such as *What is the recommended resource allocation?* and *Who is the recommended process performer?*

5   Framework

This section introduces our business-oriented framework that categorizes the identified process mining use cases. The framework can be used by practitioners to be guided in the selection of the process mining methods that are the most suitable for their needs. It consists of two parts: a categorization of the main process mining use cases (RQ1) and the elicitation of the business-oriented questions that these use cases can answer (RQ2). Our categorization draws on the value-driven business process management (VBPM) proposed in [32]. According to VBPM, organizations need BPM techniques to realize at least one of the six values: efficiency, quality, compliance, agility, integration, and networking. In order to realize these values, transparency is required. Transparency is creating visibility of how processes are executed. Commonly, this is achieved with business process models. Therefore, transparency lies at the core of VBPM.

Organizations, engaging in BPM to gain in efficiency, take an internal organizational viewpoint and focus on improving the performance of their processes. Efficiency gains are achieved by, for instance, eliminating waste in the processes, reducing redundancies, and removing rework. On the other hand, organizations can focus on the outputs of the processes and on improving their quality. Organizations that hold quality as a core value, engage in BPM to explore the correlation between process characteristics and product/service quality.

Compliance as an expected value of BPM emphasizes reducing variability and increasing standardization. Financial institutions, for instance, are subjected to regulatory requirements. Therefore, such organizations gain value from designing and executing processes that comply with predefined standards. However, organizations might also gain value from having agile processes, i.e., flexible and adaptable processes. For instance, an insurance company experiencing a sharp increase in claims during severe weather conditions, would switch to a different process execution that caters to the increased volumes. Other values of BPM are integration and networking. Integration concerns the creation of business value by increasing awareness and accessibility of process models to internal stakeholders. Conversely, networking focuses on involving external stakeholders in the processes.

We use VBPM as the basis for categorizing process mining use cases for two reasons. Firstly, VBPM is derived from surveys and interviews with companies and, therefore, captures the main reasons why companies engage with BPM. Our framework is business-oriented, i.e., categorizes use cases and questions that, when answered, aim at delivering value to organizations. Therefore, using VBPM allows us to categorize process mining methods while aligning them with the business values companies seek from BPM. Secondly, process mining is applicable in different BPM lifecycle phases such as process discovery, monitoring and analysis [25]. Likewise, process mining can be applied to support the execution of different methodologies such as Six Sigma [35]. Since VBPM focuses on business value rather than on specific BPM methodologies, using VBPM as a basis, our business-oriented framework can be used to show how process mining techniques can contribute to generating value to organizations instead of framing these techniques in the context of specific BPM lifecycle phases or methodologies.

5.1   Framework Instantiation

Our framework categorizes the main process mining use cases into categories transparency, efficiency, quality, compliance, and agility. *Transparency* encompasses use cases that aim at discovering process models (process model discovery), discovering the interaction between resources in a process (social network mining), adjusting process models to capture the process executions better (process model repair), enriching the process models with additional data (process model extension), detecting the decision rules embedded in the process decision points (rule mining), and identifying the process objectives (goal mining). *Efficiency* includes process mining use cases concerning the analysis of the performance of a process (process performance), *Quality* encompasses use cases for the identification and comparison of process variants (variant analysis) and analyzing the reasons for deviations in a process case (deviance mining). *Compliance* includes use cases comparing a process model with some observed behavior (conformance checking), or predefined rules or constraints with an ongoing case (compliance monitoring). *Agility* encompasses use cases about the predictions on how ongoing process executions will unfold in the future (predictive monitoring), the description of how the process behavior changes over time (concept drift), and the prescription of actions to take, for an ongoing case to achieve a certain desired outcome (prescriptive monitoring).

We have excluded integration and networking in our framework. Integration focuses on improving the availability and accessibility of process models to internal resources, for instance, to raise their engagement in the processes. However, although process mining methods can discover process models, the distribution and accessibility of such models are beyond the scope of this field. Networking, instead, focuses on incorporating external parties within the scope of a process. Although event logs of several parties can be merged together, process mining methods treat such logs in the same way as a single internal event log. Therefore, we also excluded networking from our framework.

Table 1. Framework instantiation.

Business-Oriented Questions Answered by Process Mining Use Cases	Sample Reference
Process Model Discovery	
How are the cases of a procedural process executed?	[43]
How are the cases of a declarative process executed?	[51]
How are the cases of a hybrid process executed?	[53]
Process Model Enhancement	
How can a process model be repaired to better reflect the actual execution of the process?	[30]
How can a process model be extended with additional information recorded in the event logs?	[64]
How can the understandability of the mined process models be improved?	[11]
Social Network Mining	
What are the relationships among the resources involved in a process?	[2]
Goal Mining	
What are the process goals?	[74]
Rule Mining	
What are the decision rules that determine which path a case takes?	[62]
Process Performance	
What is the duration-related performance of a case?	[40]
What is the resource-related performance of a case?	[39]
What is the quality-related performance of a case?	[4]
How is the performance of a case affected by other factors?	[37]
What is the performance of multiple connected processes?	[28]
How can the performance of a case be optimized?	[76]
How does the performance of a case change over time?	[21]
Process Variant Analysis	
What are the main variants of a process?	[67]
What are the similarities between two or more variants of a process?	[38]
What are the differences between two or more variants of a process?	[10]
Process Deviance Mining	
Why do some cases deviate from the normal flow?	[55]
Process Conformance Checking	
Where does a case differ from a process model?	[42]
Given a non-compliant case, what is closest compliant case?	[3]
Process Compliance Monitoring	
Is an ongoing case compliant with some predefined rules and constraints?	[52]
What are the causes of non-compliance in an ongoing case?	[9]
Predictive Monitoring	
What is the predicted remaining time of an ongoing case?	[58]
What are the predicted delays of an ongoing case?	[65]
What are the next activity/activities of an ongoing case?	[66]
What is are the predicted waiting times of an ongoing case?	[7]
What are the predicted outcomes of an ongoing case?	[47]
What are the predicted risks of an ongoing case?	[14]
What are the predicted costs of an ongoing case?	[72]
What are the predicted performance indicators of an ongoing case?	[17]
Concept Drift	
How has the process behavior changed over time?	[49]
Prescriptive Monitoring	
What is the recommended execution path of an ongoing case?	[71]
What is the recommended resource allocation?	[75]
Who is the recommended process performer?	[44]
When should interventions be made to increase the probability of a positive outcome?	[69]

The groups from top to bottom are labeled in the left margin: TRANSPARENCY, EFFICIENCY, QUALITY, COMPLIANCE, AGILITY.

At the highest level, our framework categorizes process mining use cases into transparency, efficiency, quality, compliance, and agility. Each of these categories is then organized in sub-categories. For instance, transparency consists of process model discovery, repair, enhancement, social network mining, goal mining, and rule mining (see first column of Table 1). Then, we define the questions that each use case of a certain sub-category can answer, and provide a sample reference[3] (see second column of Table 1). For instance, for concept drift under agility, we have defined the question *How has the process execution changed over time?*

The questions specified in the transparency category are, as expected, descriptive, while questions specified in the efficiency category are quantitative or comparative. The questions specified in the quality and in the compliance categories are, instead, descriptive, comparative, or explanatory. Finally, the questions defined in the agility category are descriptive, recommendatory, comparative, or explanatory. Thus, we can observe that descriptive questions lie at the foundation of other questions, just as transparency constitutes the foundation of the other categories.

5.2   Limitations

The main limitations of SLR studies are selection bias and data extraction inaccuracies. These threats, although not eliminated, were reduced by adhering to the guidelines proposed by [45]. More specifically, we used well-known databases to find papers, performed backwards referencing to avoid excluding potentially relevant papers, and ensured replicability by providing access to the SLR protocol. Another limitation concerns the fact that we relied on the results reported in the literature and we did not empirically verify or assess the extent to which the use cases impact the business processes, or if they led to effective process improvements (we considered methods using real-life event logs but not necessarily tested in industrial contexts). Although this could represent a limitation for the generalizability of the results, the proposed framework still provides practitioners with valuable insights on how process mining might be applied in industry, and represents an easy-to-use instrument to understand what types of analysis can be conducted with the existing process mining methods.

6   Conclusion

Process mining methods have been growing fast in the last decades. While such methods help manage business processes, it can be challenging for practitioners to readily understand how they can deliver value, or what business-oriented questions they can answer. To fill this research gap, we propose a framework that classifies existing process mining use cases using categories transparency, efficiency,

---

[3] Due to space limitations, only sample references are included. The complete framework can be accessed at https://doi.org/10.6084/m9.figshare.17099402.

quality, compliance, and agility. Furthermore, within each of the above categories, process mining use cases can answer descriptive, comparative, explanatory, or recommendatory questions.

The SLR we conducted also highlights that several studies in the process mining literature support the discovery of process models (transparency), predictive monitoring (agility), the analysis of process performance and variants (efficiency and quality), and conformance checking (compliance). In this respect, the framework also represent an instrument allowing researchers and/or process mining companies to understand which use cases in the process mining field have already been largely explored and which ones, instead, need further investigations.

References

1. van der Aalst, W.M.P.: Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-19345-3
2. van der Aalst, W.M.P., Reijers, H.A., Song, M.: Discovering social networks from event logs. Comput. Support. Coop. Work **14**(6), 549–593 (2005)
3. Adriansyah, A., van Dongen, B.F., van der Aalst, W.M.P.: Conformance checking using cost-based fitness analysis. In: Proceedings of the 15th IEEE International Enterprise Distributed Object Computing Conference, EDOC 2011, Helsinki, Finland, pp. 55–64. IEEE Computer Society (2011)
4. Arpasat, P., Porouhan, P., Premchaiswadi, W.: Improvement of call center customer service in a thai bank using disco fuzzy mining algorithm. In: 2015 13th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015), pp. 90–96. IEEE (2015)
5. Augusto, A., et al.: Automated discovery of process models from event logs: review and benchmark. IEEE Trans. Knowl. Data Eng. **31**(4), 686–705 (2019)
6. Batista, E., Solanas, A.: Process mining in healthcare: a systematic review. In: 9th International Conference on Information, Intelligence, Systems and Applications, IISA 2018, Zakynthos, Greece, 23–25 July 2018, pp. 1–6. IEEE Computer Society (2018)
7. Benevento, E., Aloini, D., Squicciarini, N., Dulmin, R., Mininno, V.: Queue-based features for dynamic waiting time prediction in emergency department. Meas. Bus. Excell. **23**(4), 458–471 (2019)
8. Bergami, G., Di Francescomarino, C., Ghidini, C., Maggi, F.M., Puura, J.: Exploring business process deviance with sequential and declarative patterns. CoRR abs/2111.12454 (2021)
9. Böhmer, K., Rinderle-Ma, S.: Mining association rules for anomaly detection in dynamic process runtime behavior and explaining the root cause to users. Inf. Syst. **90**, 101438 (2020)
10. Bolt, A., de Leoni, M., van der Aalst, W.M.P.: Process variant comparison: using event logs to detect differences in behavior and business rules. Inf. Syst. **74**, 53–66 (2018)

11. Bose, R.P.J.C., van der Aalst, W.M.P.: Trace clustering based on conserved patterns: towards achieving better process models. In: Rinderle-Ma, S., Sadiq, S., Leymann, F. (eds.) BPM 2009. LNBIP, vol. 43, pp. 170–181. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12186-9_16

12. Burattin, A., Maggi, F.M., Sperduti, A.: Conformance checking based on multi-perspective declarative process models. Exp. Syst. Appl. **65**, 194–211 (2016)

13. Carmona, J., van Dongen, B., Solti, A., Weidlich, M.: Conformance Checking: Relating Processes and Models. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99414-7

14. Conforti, R., de Leoni, M., La Rosa, M., van der Aalst, W.M.P.: Supporting risk-informed decisions during business process execution. In: Salinesi, C., Norrie, M.C., Pastor, Ó. (eds.) CAiSE 2013. LNCS, vol. 7908, pp. 116–132. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38709-8_8

15. Cooper, H.M.: Organizing knowledge syntheses: a taxonomy of literature reviews. Knowl. Soc. **1**(1), 104 (1988)

16. Corallo, A., Lazoi, M., Striani, F.: Process mining and industrial applications: a systematic literature review. Knowl. Process. Manag. **27**(3), 225–233 (2020)

17. Cuzzocrea, A., Folino, F., Guarascio, M., Pontieri, L.: A predictive learning framework for monitoring aggregated performance indicators over business process events. In: Proceedings of the 22nd International Database Engineering & Applications Symposium, IDEAS 2018, pp. 165–174. ACM (2018)

18. Dakic, D., Stefanovic, D., Cosic, I., Lolic, T., Medojevic, M., Katalinic, B.: Business process mining application: a literature review. In: Proceedings of the 29th DAAAM International Symposium, pp. 0866–0875 (2018)

19. De Giacomo, G., Maggi, F.M., Marrella, A., Patrizi, F.: On the disruptive effectiveness of automated planning for LTL_f-based trace alignment. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence, San Francisco, California, USA, 4–9 February 2017, pp. 3555–3561 (2017)

20. Deneckere, R., Hug, C., Khodabandelou, G., Salinesi, C.: Intentional process mining: discovering and modeling the goals behind processes using supervised learning. Int. J. Inf. Syst. Model. Des. (IJISMD) **5**(4), 22–47 (2014)

21. Denisov, V., Fahland, D., van der Aalst, W.M.P.: Unbiased, fine-grained description of processes performance from event data. In: Weske, M., Montali, M., Weber, I., vom Brocke, J. (eds.) BPM 2018. LNCS, vol. 11080, pp. 139–157. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98648-7_9

22. Di Ciccio, C., Maggi, F.M., Mendling, J.: Efficient discovery of target-branched declare constraints. Inf. Syst. **56**, 258–283 (2016)

23. Di Francescomarino, C., Dumas, M., Federici, M., Ghidini, C., Maggi, F.M., Rizzi, W.: Predictive business process monitoring framework with hyperparameter optimization. In: Nurcan, S., Soffer, P., Bajec, M., Eder, J. (eds.) CAiSE 2016. LNCS, vol. 9694, pp. 361–376. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39696-5_22

24. Di Francescomarino, C., Ghidini, C., Maggi, F.M., Milani, F.: Predictive process monitoring methods: which one suits me best? In: Weske, M., Montali, M., Weber, I., vom Brocke, J. (eds.) BPM 2018. LNCS, vol. 11080, pp. 462–479. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98648-7_27

25. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management, 2nd edn. Springer, Berlin (2018). https://doi.org/10.1007/978-3-662-56509-4

26. Dunzer, S., Stierle, M., Matzner, M., Baier, S.: Conformance checking: a state-of-the-art literature review. In: Proceedings of the 11th International Conference on Subject-Oriented Business Process Management, S-BPM ONE 2019, Seville, Spain, 26–28 June 2019, pp. 4:1–4:10. ACM (2019)

27. Eggers, J., Hein, A.: Turning big data into value: a literature review on business value realization from process mining. In: 28th European Conference on Information Systems, ECIS 2020 (2020)

28. Engel, R.: Analyzing inter-organizational business processes - process mining and business performance analysis using electronic data interchange messages. Inf. Syst. E Bus. Manag. **14**(3), 577–612 (2016)

29. Erdogan, T., Tarhan, A.: Systematic mapping of process mining studies in healthcare. IEEE Access **6**, 24543–24567 (2018)

30. Fahland, D., van der Aalst, W.M.P.: Model repair - aligning process models to reality. Inf. Syst. **47**, 220–243 (2015)

31. Fink, A.: Conducting Research Literature Reviews: From the Internet to Paper. Sage Publications (2019)

32. Franz, P., Kirchmer, M.: Value-Driven Business Process Management: The Value-Switch for Lasting Competitive Advantage. McGraw Hill Professional (2012)

33. Ghasemi, M., Amyot, D.: Process mining in healthcare: a systematised literature review. Int. J. Electron. Heal. **9**(1), 60–88 (2016)

34. Ghazal, M.A., Ibrahim, O., Salama, M.A.: Educational process mining: a systematic literature review. In: 2017 European Conference on Electrical Engineering and Computer Science (EECS), pp. 198–203. IEEE (2017)

35. Graafmans, T., Turetken, O., Poppelaars, H., Fahland, D.: Process mining for six sigma: a guideline and tool support. Bus. Inf. Syst. Eng. **63**(3), 277–300 (2021). https://doi.org/10.1007/s12599-020-00649-w

36. Grüger, J., Bergmann, R., Kazik, Y., Kuhn, M.: Process mining for case acquisition in oncology: a systematic literature review. In: Proceedings of the Conference on "Lernen, Wissen, Daten, Analysen", Online, 9–11 September 2020, vol. 2738, pp. 162–173. CEUR Workshop Proceedings. CEUR-WS.org (2020)

37. Hompes, B.F.A., Maaradji, A., La Rosa, M., Dumas, M., Buijs, J.C.A.M., van der Aalst, W.M.P.: Discovering causal factors explaining business process performance variation. In: Dubois, E., Pohl, K. (eds.) CAiSE 2017. LNCS, vol. 10253, pp. 177–192. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59536-8_12

38. Huang, Z., Dong, W., Duan, H., Li, H.: Similarity measure between patient traces for clinical pathway analysis: problem, method, and applications. IEEE J. Biomed. Health Inf. **18**(1), 4–14 (2014)

39. Huang, Z., Lu, X., Duan, H.: Resource behavior measure and application in business process management. Exp. Syst. Appl. **39**(7), 6458–6468 (2012)

40. Jaisook, P., Premchaiswadi, W.: Time performance analysis of medical treatment processes by using disco. In: 2015 13th International Conference on ICT and Knowledge Engineering, ICT & Knowledge Engineering 2015, pp. 110–115. IEEE (2015)

41. Jokonowo, B., Claes, J., Sarno, R., Rochimah, S.: Process mining in supply chains: a systematic literature review. Int. J. Electr. Comput. Eng. (IJECE) **8**(6), 4626–4636 (2018)

42. Kalenkova, A.A., Ageev, A.A., Lomazova, I.A., van der Aalst, W.M.P.: E-government services: comparing real and expected user behavior. In: Teniente, E., Weidlich, M. (eds.) BPM 2017. LNBIP, vol. 308, pp. 484–496. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-74030-0_38

43. Kalenkova, A.A., Burattin, A., de Leoni, M., van der Aalst, W.M.P., Sperduti, A.: Discovering high-level BPMN process models from event data. Bus. Process. Manag. J. **25**(5), 995–1019 (2019)

44. Kim, A., Obregon, J., Jung, J.-Y.: Constructing decision trees from process logs for performer recommendation. In: Lohmann, N., Song, M., Wohed, P. (eds.) BPM 2013. LNBIP, vol. 171, pp. 224–236. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06257-0_18

45. Kitchenham, B.: Procedures for performing systematic reviews. Keele University, Keele, UK **33**(2004), 1–26 (2004)

46. de Leoni, M., Marrella, A.: Aligning real process executions and prescriptive process models through automated planning. Exp. Syst. Appl. **82**, 162–183 (2017)

47. Leontjeva, A., Conforti, R., Di Francescomarino, C., Dumas, M., Maggi, F.M.: Complex symbolic sequence encodings for predictive monitoring of business processes. In: Motahari-Nezhad, H.R., Recker, J., Weidlich, M. (eds.) BPM 2015. LNCS, vol. 9253, pp. 297–313. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23063-4_21

48. Ly, L.T., Maggi, F.M., Montali, M., Rinderle-Ma, S., van der Aalst, W.M.P.: Compliance monitoring in business processes: functionalities, application, and tool-support. Inf. Syst. **54**, 209–234 (2015)

49. Maaradji, A., Dumas, M., La Rosa, M., Ostovar, A.: Detecting sudden and gradual drifts in business processes from execution traces. IEEE Trans. Knowl. Data Eng. **29**(10), 2140–2154 (2017)

50. Maggi, F.M., Corapi, D., Russo, A., Lupu, E., Visaggio, G.: Revising process models through inductive learning. In: zur Muehlen, M., Su, J. (eds.) BPM 2010. LNBIP, vol. 66, pp. 182–193. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-20511-8_16

51. Maggi, F.M., Di Ciccio, C., Di Francescomarino, C., Kala, T.: Parallel algorithms for the automated discovery of declarative process models. Inf. Syst. **74**, 136–152 (2018)

52. Maggi, F.M., Montali, M., van der Aalst, W.M.P.: An operational decision support framework for monitoring business constraints. In: 15th International Conference on Fundamental Approaches to Software Engineering, FASE 2012, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2012, Tallinn, Estonia, pp. 146–162 (2012)

53. Maggi, F.M., Slaats, T., Reijers, H.A.: The automated discovery of hybrid processes. In: Sadiq, S., Soffer, P., Völzer, H. (eds.) BPM 2014. LNCS, vol. 8659, pp. 392–399. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10172-9_27

54. Maita, A.R.C., Martins, L.C., Paz, C.R.L., Rafferty, L., Hung, P.C.K., Peres, S.M., Fantinato, M.: A systematic mapping study of process mining. Enterp. Inf. Syst. **12**(5), 505–549 (2018)

55. Mannhardt, F., de Leoni, M., Reijers, H.A., van der Aalst, W.M.P.: Data-driven process discovery - revealing conditional infrequent behavior from event logs. In: Dubois, E., Pohl, K. (eds.) CAiSE 2017. LNCS, vol. 10253, pp. 545–560. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59536-8_34

56. Milani, F., Maggi, F.M.: A comparative evaluation of log-based process performance analysis techniques. In: Abramowicz, W., Paschke, A. (eds.) BIS 2018. LNBIP, vol. 320, pp. 371–383. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93931-5_27

57. Naderifar, V., Sahran, S., Shukur, Z.: A review on conformance checking technique for the evaluation of process mining algorithms. TEM J. **8**(4), 1232 (2019)

58. Navarin, N., Vincenzi, B., Polato, M., Sperduti, A.: LSTM networks for data-aware remaining time prediction of business process instances. In: 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017, Honolulu, HI, USA, pp. 1–7. IEEE (2017)
59. Nuritha, I., Mahendrawathi, E.: Behavioural similarity measurement of business process model to compare process discovery algorithms performance in dealing with noisy event log. Procedia Comput. Sci. **161**, 984–993 (2019)
60. Okoli, C.: A guide to conducting a standalone systematic literature review. Commun. Assoc. Inf. Syst. **37**, 43 (2015)
61. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: a literature review. J. Biomed. Inf. **61**, 224–236 (2016)
62. Rozinat, A., van der Aalst, W.M.P.: Decision mining in ProM. In: Dustdar, S., Fiadeiro, J.L., Sheth, A.P. (eds.) BPM 2006. LNCS, vol. 4102, pp. 420–425. Springer, Heidelberg (2006). https://doi.org/10.1007/11841760_33
63. dos Santos Garcia, C., et al.: Process mining techniques and applications - a systematic mapping study. Exp. Syst. Appl. **133**, 260–295 (2019)
64. Seeliger, A., Stein, M., Mühlhäuser, M.: Can we find better process models? Process model improvement using motif-based graph adaptation. In: Teniente, E., Weidlich, M. (eds.) BPM 2017. LNBIP, vol. 308, pp. 230–242. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-74030-0_17
65. Senderovich, A., Weidlich, M., Gal, A., Mandelbaum, A.: Queue mining for delay prediction in multi-class service processes. Inf. Syst. **53**, 278–295 (2015)
66. Tax, N., Verenich, I., La Rosa, M., Dumas, M.: Predictive business process monitoring with LSTM neural networks. In: Dubois, E., Pohl, K. (eds.) CAiSE 2017. LNCS, vol. 10253, pp. 477–492. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59536-8_30
67. Taymouri, F., La Rosa, M., Dumas, M., Maggi, F.M.: Business process variant analysis: survey and classification. Knowl. Based Syst. **211**, 106557 (2021)
68. Teinemaa, I., Dumas, M., La Rosa, M., Maggi, F.M.: Outcome-oriented predictive process monitoring: review and benchmark. ACM Trans. Knowl. Discov. Data **13**(2), 17:1–17:57 (2019)
69. Teinemaa, I., Tax, N., de Leoni, M., Dumas, M., Maggi, F.M.: Alarm-based prescriptive process monitoring. In: Weske, M., Montali, M., Weber, I., vom Brocke, J. (eds.) BPM 2018. LNBIP, vol. 329, pp. 91–107. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98651-7_6
70. Thiede, M., Fuerstenau, D., Barquet, A.P.B.: How is process mining technology used by organizations? A systematic literature review of empirical studies. Bus. Process. Manag. J. **24**(4), 900–922 (2018)
71. Thomas, L., Kumar, M.M., Annappa, B.: Recommending an alternative path of execution using an online decision support system. In: Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence, pp. 108–112 (2017)
72. Tu, T.B.H., Song, M.: Analysis and prediction cost of manufacturing process based on process mining. In: 2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA), pp. 1–5. IEEE (2016)
73. Verenich, I., Dumas, M., La Rosa, M., Maggi, F.M., Teinemaa, I.: Survey and cross-benchmark comparison of remaining time prediction methods in business process monitoring. ACM Trans. Intell. Syst. Technol. **10**(4), 34:1–34:34 (2019)
74. Yan, J., Hu, D., Liao, S.S.Y., Wang, H.: Mining agents' goals in agent-oriented business processes. ACM Trans. Manag. Inf. Syst. **5**(4), 20:1–20:22 (2015)

75. Zhao, W., Liu, H., Dai, W., Ma, J.: An entropy-based clustering ensemble method to support resource allocation in business process management. Knowl. Inf. Syst. **48**(2), 305–330 (2016)
76. Zhao, W., Yang, L., Liu, H., Wu, R.: The optimization of resource allocation based on process mining. In: Huang, D.-S., Han, K. (eds.) ICIC 2015. LNCS (LNAI), vol. 9227, pp. 341–353. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22053-6_38