2023 International Conference on Frontiers of Information Technology (FIT) | 979-8-3503-9578-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/FIT60620.2023.00018

Survey of Explainability within Process Mining: A case study of BPI challenge 2020

Tjalling Hoogendoorn
Dept. of Computer Science
University of Twente
Enschede, the Netherlands
t.a.d.hoogendoorn@student.utwente.nl

Jeewanie Jayasinghe Arachchige
Dept. of Computer Science
University of Twente
Enschede, the Netherlands
j.a.jayasinghearachchige@utwente.nl

Faiza A. Bukhsh
Dept. of Computer Science
University of Twente
Enschede, the Netherlands
f.a.bukhsh@utwente.nl

***Abstract*—The need for explainability in Business process management is tremendously increasing, especially in the age of generative AI. The number of published articles on explainable AI (XAI) has skyrocketed for five years. AI impacts the decision-making process in business analytics. Process mining as a sub-discipline of data science can play a role in explainable business decision-making. Process mining exhibits its intention in process discovery, performance measures of processes, and process improvements based on the event logs. Although the accuracy of the outcome of process mining models has been investigated at a certain level, the explainability of those is possible through the discretization of the analytic steps. As an initial step in exploring the explainability of process mining, this research conducts a technical analysis of 37 research papers submitted to the Business Process Intelligence (BPI) Challenge 2020. The main focus of this analysis aims to answer the question, "How and why a process model is produced? " To make a foundation for the research question, the notion of explainability is explored based on an Explainable AI ontology. Due to the small sample size, the study cannot identify clear trends of explainability in process mining. However, the results conclude that explainability depends on the process model's transparency and reproducibility. Moreover, further research with a large sample size is required to understand the discrete factors impacting decision-making in business process management.**

***Index Terms*—Data Science, Process Mining, Explainability, Transparency**

I. Introduction

A large quantity of data is produced as software has been able to log and store data [1] automatically. Moreover, using the software logs, computers can now navigate the flow of time, analyzing, dissecting, and visualizing events from completely new angles. [2]. Process Mining (PM) is an emerging discipline that analyze event logs and produces process models with different levels of complexities [1]. Moreover, process mining identifies bottlenecks and anomalies within the process [1], [3], [4]. However, there are many challenges while visualizing the complex processes for analysis, which often lead to unexplained models which are difficult to understand without context and background information. Moreover, the robust interpretation of process deviations is a crucial challenge during the application of process mining in diverse domains such as healthcare and logistics [5], [6]. Thus, there is a need for understandable models. The capability of being understandable is called explainability. The problem described above can be solved with the help of explainable process models, which take complex data and provide a simple solution with the relevant explanation for robust interpretation.

There is a dearth of robust interpretation while analyzing event logs. During the analysis of temporal data, different levels of filtering, mapping and rendering are applied to achieve different views of the same process [2]. Thus, various PM algorithms and their configurations provide different views of the similar data set [3]. The best model to provide insight into a certain process is the one that strikes a balance between fitness (how well the model allows for the events in the log), simplicity, generalization and precision (not over-fitting or under-fitting) [3]. Particularly, during process discovery, it is important to have a visual that uncovers the patterns of the given data. However, due to the variation in processes, it is up to the analyst to find the best tool for the job, and process mining techniques appear to provide new possibilities.

Recently, PM as an emerging discipline has been adopted by non-practitioners [7]. Therefore, the explainability of PM methods/algorithms is very important. Prior to exploring the explainability of PM, it is worthwhile to discover the essence of the term explainability itself. Thus, the main objective of this research work is as as follows; (As an extended work of the first author's thesis [8])

- To investigate the prevalence of explainability in process mining techniques and produced models.

The rest of the paper is structured as follows, Section two discusses the background information. Section three describes the notion of explainability and its adaptability in process mining. Further, sections four and five present the research approach and results, respectively. At last, section six concludes the paper.

II. Background Information

In order to conduct this research, a data set is applied that is published by the IEEE Process Mining Task Force (PMTF) during the Business Process Intelligence (BPI) Challenge [9]. The PMTF aims to promote research in the adoption of process mining in the academic and professional community. Thus, the PMTF arranges the BPI Challenge and invites the processed data from the given data sets. The application of

2473-7569/23/$31.00 ©2023 IEEE
DOI 10.1109/FIT60620.2023.00018

PM algorithms on these data sets explains a complex process in a single context in.

In literature, PM has been discussed in general. However, there is a dearth of focused research in this field with respect to explainability. Due to the nature of the data, there will be a limited focus on analyzing data traces from an event log after a process has finished (considered a post-mortem), which stands in contrast to analyzing a process while it is still happening (pre-mortem). Furthermore, the study focuses on process discovery and process compliance enhancement. More information is provided in section II.C in [3].

In general, the topic of explainability is elaborated with regard to Artificial Intelligence (AI). The literature provides the conceptual framework of explainable process prediction in AI in [10] depicted in Figure 1.

Although explainability for deep learning and other black-box approaches have already been studied in other domains, there are a few explainable approaches in the process mining domain [10]. The AI is a black box that makes decisions, and the transparency of decision-making is unknown. Explainability is the art of making the decision process transparent and giving context for a human to interpret the decision made by AI [11]. Thus, explainability plays a vital role in the acceptance of black box models, such as AI models, by adding transparency to the decision-making process. However, AI does not only face the challenge of transparency but the upsurge in advanced Machine Learning models, powered by large amounts of gathered and curated data, and data science as a whole, faces these challenges of transparency and correct use of data. Process mining predicts and analyzes processes and steers them in a certain way, and also requires the responsible use of the data [7]. Therefore, the transparent interpretation of the application of process mining is a crucial challenge within the field of PM.

III. Notion of Explainability

The necessity of explainability within PM is equally important as the accuracy of PM models. However, the notion of explainability within PM is not concretely defined yet. Therefore, first, we look into the concept of explainability within PM. There are numerous types of research that appeared in the literature that focused on the accuracy of PM models but explainability is not discussed. However, the systematic literature review done in the area of XAI by Giulia Vilone and Luca Longo [12] reveals that the notion of explainability is related to interpretability, understandability, comprehensibility, and justifiability. Further, it stated various constructs of explainability might be thought of as a concept borrowed from psychology since it is strictly connected to humans. These concepts are similarly valid to explainability within PM as well. However, the authors concluded that there is still no agreement among scholars regarding the notation of explainability in AI. Further, the important properties that should be considered to make it effective and understandable for end-users have not finalized so far in XAI. In our general view, the explainability of PM allows the PM approach more transparent and provides better understandability about the decision to the end users. We strongly agree that future research is required to identify the prominent constructs for the explainability of PM. Thus, prior to carrying out the survey, it is important to investigate the meaning of ”explainability” within the PM domain. The following section aims to establish a foundation for understanding the concept of ”explainability.

A. *Ontology for Taxonomy in XPM*

As the explainability of process mining does not appear in the literature, we tried to guide our research with the research field XAI based on several commonalities. Both AI and PM are centered around algorithmic decisions, and both follow a data-driven approach. Therefore, this research adopted the taxonomy provided for XAI and its high-level ontological representation proposed by [13] with respective alternations to define explainability in PM. Figure 2 exhibits the ontology for the taxonomy for explainability adopted to PM. The ontology proposed for the explainability in [13] has two paths to achieve ”Explainable Models”. One path is very direct that is from ”Transparent Models” in which the underline algorithm has high transparency. However, the opaque models require post hoc treatments to make them explainable. When translating the “Opaque Model” towards to explainable model, [13] suggested post hoc treatments. These post hoc treatments were classified into two categories that model model-specific and model-agnostic. The first one refers to the explainable techniques that are strictly dependent on the model. However, we consider only the agnostic model, as the PM models do not depend on any platform, operating system, etc. Generic techniques used to enhance the explainability of “Opaque Models,” as listed under model-agnostic approaches, are valid explainable techniques for PM as well. Therefore, we refined the Ontology and taxonomy with PM-specific algorithms and features.

The taxonomy is described in the following subsection and ontology is exhibited in Figure 2.

- Transparent model: Transparent models include Alpha miner, heuristics and inductive miner. The outputs from these models are often transparent [14]. However, the level of explainability is not explicit.
- Opaque model: Opaque models are generated from fuzzy miner, genetic miner and spectrum miner [15].
- Explanation by simplification: By simplifying a model via approximation [16] and abstraction, we can find alternatives to the original models to explain the prediction we are interested in.
- Explanation by feature relevance: This evaluates features such as fitness, simplicity, generalization and precision [5] based on its average expected marginal contribution to the model’s decision,
- Visual explanation: Process visualization techniques used to interpret the outcome of PM [17].
- Local explanation: Approximate the model into a narrow area around a specific instance of interest, and offer information about how the process model operates when

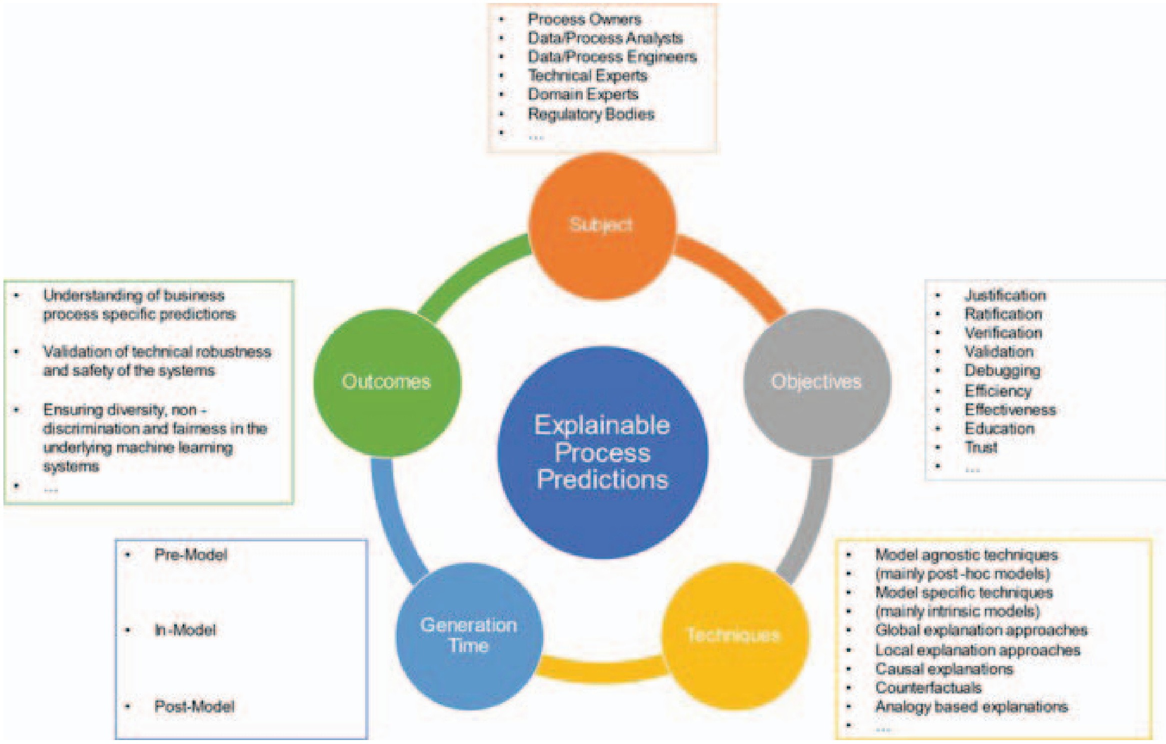

Fig. 1: A conceptual framework for explainable process prediction [10]

encountering inputs that are similar to the one we are interested in explaining.

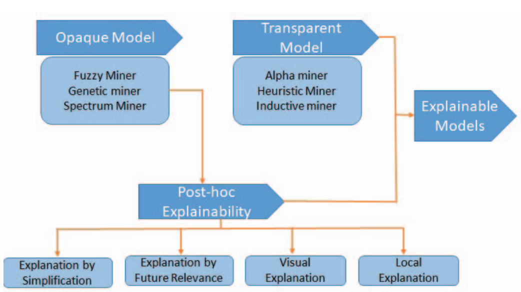

Fig. 2: Ontology of Explainable PM taxonomy adapted from [13]

IV. Research Approach

The study adopts PM^2 model for evaluation that is described by van der Aalst et al [18]. Moreover, the study follows the iterative approach (2019) [19] as the methodology because the current study conducts the literature review of process mining. Thiede et al. [20] and Ghasemi and Amyot [21] provided some rigorous approaches such as PRISMA [22] that have applied in the medical literature reviews. Alternatively, another approach that is presented by Snyder [23], has not been adopted by this study, due to difference of scope of the research. There is a dearth of robust literature review methodology in the field of process mining. The study has selected a semi-systematic general literature review as mentioned by Garcia et al. [19].

A. Approach on Data Collection

BPI Challenge data has often been reused for research in the past [20]. We selected BPI2020 dataset [1], and an evaluation was carried out for all 37 papers submitted to this challenge. Each paper is evaluated on the following conditions:

- Which platform is used?
- Which techniques are used to create the process models?

Then the evaluated results are summarised in a table, focusing exclusively on papers that (mainly) use ProM, an open-source mining tool of which we can see how algorithms are implemented (popular proprietary alternatives are Fluxicon Disco and Celonis, all of them are closed-source tools).

B. Approach on Data Processing

After all the data was gathered, it was categorized based on platform usage which main platform was used and how often each platform is used. From observation of the gathered data, we distinguish the following: For the main platform, we find the following unique categories with regard to what was used for PM analysis:

- Celonis - Uses Celonis exclusively for PM techniques
- Disco - Uses Disco for PM techniques exclusively

[1]Full list found at icpmconference.org/2020/bpi-challenge/

- ProM - Uses ProM exclusively for PM techniques
- PM4Py - Uses exclusively PM4Py library and Python scripts for PM techniques
- Mixed - Used a combination of the above 4 methods for different PM analysis techniques and/or state multiple platforms but not which analysis was performed with which platform.
- Other - Used own software or did not use PM techniques

Figure 3 shows the results of this observation.

Furthermore, the following unique categories are defined for frequency analysis: Celonis, Disco, ProM, PM4Py, Non-PM methods and other (own software, other libraries) The results of this observation is in Figure 4.

V. Results

A total of 37 papers have been submitted for the ICMP conference BPI Challenges 2020. Three papers of this collection did not use modern process mining techniques, constructing BPM or Petri net graphs from raw statistical data instead. Seven papers have a platform but no explanation of how the researchers derived these results and 11 papers have a platform but do not state the algorithms used. Out of the last 16 papers, 13 mentioned an ”ideal process graph” or process visualization/throughput analysis. Table I shows the three papers which used ProM (primarily) and explicitly stated which techniques were used for process discovery.

Paper Number	Platform	Techniques Used
126	ProM	Inductive, Heuristic, Fuzzy Miner
146	Disco, ProM, PM4Py	Inductive, Performance spectrum Miner
150	ProM	Inductive Visual Miner

TABLE I: Three papers having a clear technique description and using open source software.

A. Observations on Platform Use

From our dataset, we can observe how often each platform is used and which platforms are most prevalent as a main platform in figures 4 and 3. The difference here is that certain papers use an equal mix of different platforms for different analyses (the total use of which is analyzed in figure 4). However, some papers clearly favor one platform for the majority of their analysis, which is visualized in figure 3. Figure 3 shows that Celonis is most often used as a primary or sole platform in their research. However, we observe that a lot of practitioners use a mix of two or more platform. The histogram in figure 4 shows that Disco is the most frequently used platform, with Celonis and PM4py being second.

B. Result Analysis and explanabaility

In this section, the results derived in the previous subsection is analyzed with focusing to the explainability within PM. The first important observation is that we use 37 papers, while an average literature study uses hundreds. As such the power of our conclusions here is minimal.

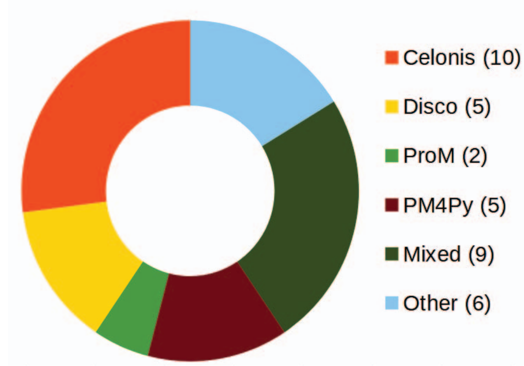

Fig. 3: Donut chart showing which single platforms are prefered vs an equal mix of platforms of BPI 2020 papers

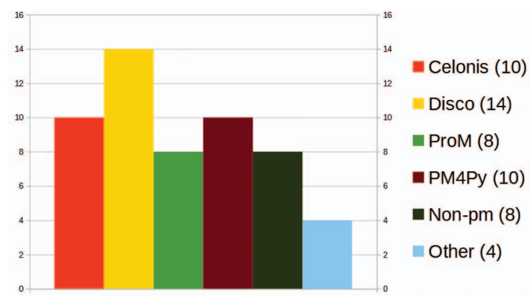

Fig. 4: Frequency histogram of total use of platforms BPI 2020 papers

The observation derived from figure 3 reveals that there is a strong preference for Celonis (a closed-source platform) for all forms of analysis. This value may have been influenced by several papers which explicitly credited Celonis Educational for the use of their platform for their course. Thus this value is likely skewed. Disregarding Celonis, we observe that most participants prefer a mixed approach where multiple platforms are used for analysis. We find the open source platforms ProM and PM4Py to be roughly equal with closed source competitor Disco.

The frequency analysis of the total usage of platforms in figure 4 shows that the most popular platform is actually Disco (including mixed-use). Disco Visual analysis pops up in all mixed platforms (while ProM is able to provide the same function, Disco is favored here). It shows a trend to use specifically Disco’s Visual Miner for early data exploration.

According to the findings which described at the beginning of section 5, only 3 papers out of 37 mentioned which specific miner has been used. This value is likely influenced by the fact that only PM4Py and ProM require you to think about specific miners at all. Celonis and Disco abstract part of this process away in favor of more user-friendliness. Thus the transparency of how specifically an ideal process model is generated, and what happens ”under the hood” is simply not visible depending on your choice of platform.

Since we find that the difference between open and closed source platforms appears to matter for how transparent a writer can be, it is important to look at the ratio of open and closed source platforms. Based on figure 4, it can be found a total of 24 closed platform uses in Celonis and Disco. The open platforms PM4Py and ProM constitute 18 counts. On top of that, we can add the other category, as it constitutes other open-source libraries and hobbyists making their own PM analysis of which they published the code. Non-PM methodology is not considered for this analysis. The end result is 24 closed platforms vs 22 open-source platforms, which means that its use is spread rather equally over the sample space. However, figure 3 shows that open-source platforms are rarely used as a primary platform for the bulk of the analysis. Thus we can conclude that closed-source platforms are slightly more popular, due to their ease of use.

Finally, when analyzing the trends for specific process mining algorithms since only 3 papers mentioned their techniques explicitly, we conclude that there is not enough data to indicate a clear trend. However, we can conclude that the use of visual miners for early analysis is very popular.

The categorization of papers was done in a semi-rigorous fashion. Some definitions have not been clearly defined or marked in the data. Additional insights could have been gathered but have not been considered due to lack of time.

For this research, no distinction was made between professional papers and student entries, which means it is unclear in this study if there is a difference between these fields that can have influenced the study. However, there is a strong variety in the quality of the papers.

Due to the lack of information, it is unclear if the lack of explanation is indicative of author's inability to explain or the tools/algorithms limitation. Explainable AI taxonomy mentioned in section **??** helps to understand the reasoning for lack of transparency for reproducibility.

C. Limitations

This research has a narrow focus compared to typical literature reviews, covering only 37 papers. Furthermore, because of the choice of source, the 2020 BPI Challenge, there is a mix of students who participated in the challenge as part of a university course, professionals who participated because their company mandates it, and those who have an intrinsic interest in the BPI challenge. As such there might be a bias for lower-quality reporting due to the nature of the participants. Similarly, because of the mixed nature, it is unclear how representative the data is for any one demographic.

This study, due to the narrow scope of papers that it reviewed, cannot provide justified insights into the quality of PM reporting by practitioners in the field but can be considered as a first milestone toward explainable process mining.

D. Future Research

During the course of this research, the following avenues for potential follow-up research have been identified:

- Broadening the scope of papers covered in the review and understanding the explainability aspects in PM.
- Making more fine-grained meta-observations on professional vs student entries, explaining data discovery, and Interpreting the process model.
- Repeat this study but with a more selective sample demographic.
- Perform Social research on how much explanation is necessary for a non-practitioner to feel they understand the used methodology.
- Perform a follow-up integrative review to see how lessons from psychology and/or Explainable AI can be applied in this field.

VI. Conclusion

This research paper aims to analyze explainability within process mining, to get better insights. The focus is on the future conversation about how to explain Process Mining as a science to a wider audience. During the course of this paper, it was found that the analyzed sample provided inconclusive insights.

Significant conclusions from this paper are that the Explanability aspect is directly relevant to phase 3 (data processing) and phase 4 (mining and analysis) of PM2 methodology. All other phases of PM2 can be considered as not influenced by explainable AL taxonomy. Moreover, in the BPI Challenge sample, it is highly uncommon to state exactly what technology is used ”under the hood” to generate a process model, either due to this information being abstracted away from the user or not mentioned in the paper. Closed source platforms abstract this knowledge away from the user, but due to the distribution of 24 closed source vs 22 open source platforms used, this can not be the sole explaining factor. However, this statistic indicates that there appears to be no clear trend with regard to open or closed sources. There is, however, a trend within this sample, not to mention how process models are generated. Further research in a systematic fashion is necessary to determine if this is just because of the focus of the challenge not being on explainability or if this trend also exists in the larger body of process mining literature.

VII. Acknowledgement

The authors would like to thank Professor Y.J. Erden of the Faculty of Philosophy at the University of Twente for her valuable discussion. Further, the authors would like to thank dr. Usman Ahmad of the faculty of Behavioral, Management, and Social Sciences for his contribution.

References

[1] W. van der Aalst, *Process Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.

[2] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of Time-Oriented Data*. Human-Computer Interaction Series, London: Springer London, 2011.

[3] N. Chaydy and A. Madani, "An overview of Process Mining and its applicability to complex, real-life scenarios," in *2019 International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoBIoTS)*, (Casablanca, Morocco), pp. 1–9, IEEE, Dec. 2019.

[4] S. Miksch, "Visual Analytics Meets Process Mining: Challenges and Opportunities," in *2021 3rd International Conference on Process Mining (ICPM)*, (Eindhoven, Netherlands), pp. xiv–xiv, IEEE, Oct. 2021.
[5] A. Yeshchenko, C. Di Ciccio, J. Mendling, and A. Polyvyanyy, "Visual Drift Detection for Event Sequence Data of Business Processes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, pp. 3050–3068, Aug. 2022.
[6] M. Noshad, C. C. Rose, and J. H. Chen, "Signal from the noise: A mixed graphical and quantitative process mining approach to evaluate care pathways applied to emergency stroke care," *Journal of Biomedical Informatics*, vol. 127, p. 104004, Mar. 2022.
[7] W. M. P. van der Aalst, "Responsible Data Science: Using Event Data in a "People Friendly" Manner," in *Enterprise Information Systems* (S. Hammoudi, L. A. Maciaszek, M. M. Missikoff, O. Camp, and J. Cordeiro, eds.), vol. 291, pp. 3–28, Cham: Springer International Publishing, 2017.
[8] T. Hoogendoorn, "Survey of explainability within process mining," B.S. thesis, University of Twente, 2023.
[9] TFPM, "BPI Challenge – Process Mining Conference 2020," 2020.
[10] N. Mehdiyev and P. Fettke, "Explainable artificial intelligence for process mining: A general overview and application of a novel local explanation approach for predictive process monitoring," *Interpretable Artificial Intelligence: A Perspective of Granular Computing*, pp. 1–28, 2021.
[11] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining Explanations: An Overview of Interpretability of Machine Learning," in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, (Turin, Italy), pp. 80–89, IEEE, Oct. 2018.
[12] V. Giulia, , and L. Longo, "Notions of explainability and evaluation approaches for explainable artificial intelligence," *Journal of Information Fusion*, vol. 76, pp. 89–106, May 2021.
[13] P. P. Angelov, E. A. Soares, R. Jiang, N. I. Arnold, and P. M. Atkinson, "Explainable artificial intelligence: an analytical review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 11, no. 5, p. e1424, 2021.
[14] H. S. Bernis Tibeme and C. Zhang, "Process mining algorithms for clinical workflow analysis," IEEE, May 2018.
[15] E. P. Gupta, "Process mining a comparative study," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, May 2014.
[16] J. Tritscher, M. Ring, D. Schlr, L. Hettinger, and A. Hotho, *Evaluation of Post-hoc XAI Approaches Through Synthetic Tabular Data*, pp. 422–430. 09 2020.
[17] B. F. van Dongen and A. Adriansyah, *Process Mining: Fuzzy Clustering and Performance Visualization*, pp. 159–169. 2009.
[18] M. L. van Eck, X. Lu, S. J. J. Leemans, and W. M. P. van der Aalst, "PM$$\hat{2}$$: A Process Mining Project Methodology," in *Advanced Information Systems Engineering* (J. Zdravkovic, M. Kirikova, and P. Johannesson, eds.), vol. 9097, pp. 297–313, Cham: Springer International Publishing, 2015.
[19] C. d. S. Garcia, A. Meincheim, E. R. Faria Junior, M. R. Dallagassa, D. M. V. Sato, D. R. Carvalho, E. A. P. Santos, and E. E. Scalabrin, "Process mining techniques and applications – A systematic mapping study," *Expert Systems with Applications*, vol. 133, pp. 260–295, Nov. 2019.
[20] M. Thiede, D. Fuerstenau, and A. P. Bezerra Barquet, "How is process mining technology used by organizations? A systematic literature review of empirical studies," *Business Process Management Journal*, vol. 24, pp. 900–922, Jan. 2018.
[21] M. Ghasemi and D. Amyot, "Process mining in healthcare: A systematised literature review," *International Journal of Electronic Healthcare*, vol. 9, no. 1, p. 60, 2016.
[22] D. Moher, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," *Annals of Internal Medicine*, vol. 151, p. 264, Aug. 2009.
[23] H. Snyder, "Literature review as a research methodology: An overview and guidelines," *Journal of Business Research*, vol. 104, pp. 333–339, Nov. 2019.