

How Do Explainability and Interpretability Affect the Use of Process Mining?

Irina Tentina¹

¹The Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

Process mining continues to mature from a technological perspective. However, despite the increasing number of process mining solutions and services, some barriers still remain to its full adoption in business practice. Like any other technology, process mining relies on effective usage by users to add value. One of the factors affecting the adoption by users and the use of process mining is that users cannot understand and translate process mining outputs into valuable insights. Our hypothesis is that interpretation and explainability are crucial to give users a sense of understanding and actionability. This doctoral project aims to define explainability and interpretability in process mining, investigate different factors affecting them and, furthermore, design solutions for improvement. The project follows a mix of empirical and technical approaches. As a pilot study survey-based interviews with PM business users were conducted confirming the relevance of the topic.

Keywords

Process Mining, Explainability, Interpretability

1. Introduction and Motivation

Process mining (PM) is a technology that uses event logs from IT systems and applications to reconstruct, visualize, analyze and improve business processes [1]. According to Gartner [2] the PM field continues to grow and the PM software market grew by 39.5% to \$871.6 million in 2023. The number of PM solutions and tools is increasing every year. Most large market players, such as SAP, Microsoft and IBM, acquire smaller PM companies to include PM in their solution stack. There has also been a growth of consulting companies offering different types of PM services.

Both the increase in solutions and consultancy seem to be indicative of the demand for PM among companies within the industry. At the same time, adoption by end users still appears to be an issue. There are more companies that either already had some experience with PM solution(s) and decided not to continue with it [3] or still use it, but struggle with adoption, value identification and realization [4]. There are many factors that can lead to (un-)successful PM use. In this research project we focus on challenges, as it is called below - **pain points** (PP), which stem from the author's consulting experience and have been highlighted in the recent research[5].

PP1: General PM eXplainability (XPM) - I do not always understand what exactly am I seeing. PM outcomes can be presented in different ways and, depending on the PM solution, may include process-focused visualizations such as process maps and/or set of different visual components as dashboards. However, users who are not experts in PM may experience reading these visuals and understanding exactly what is displayed as "overwhelming" and non-trivial[6].

PP2: General PM iNterpretability (NPM) - It is challenging to understand what should I do with what I see. Once users understand what they see, e.g. how to read a process map and/or dashboard, the next step is to interpret what they see and translate it into valuable insights. Depending on the project setting, different approaches and techniques can be used to analyze PM results, for example exploratory and/or confirmatory approach[7]. However, when non-expert users start using one of these techniques, they may still find it challenging to interpret PM results properly without help from PM experts[6].

ICPM 2024 Doctoral Consortium, October 14–18, 2024, Kongens Lyngby, Denmark

✉ i.tentina@tue.nl (I. Tentina)

🆔 0009-0006-2281-1201 (I. Tentina)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

PP3: PM without Process Focus - I miss Process Mining in all this. Because of PP1 and PP2, PM developers may try to avoid complex output and simplify analyses to improve NPM. In some cases, this can lead to better operational use, for example, when a non-expert user is given only a short list of tasks to focus on. However, it can also lead to using PM without a process focus or with a focus only on local questions within the complex end-to-end view, similar to Business Intelligence (BI) tools. The question then arises as to whether such simplification helps to improve XPM and NPM, or whether it risks that actual complex problems remain unsolved.

2. Research Questions

To investigate pain points mentioned in the previous section, we have formulated the following research questions.

RQ1: How can XPM and NPM be defined for certain stages of the PM journey for certain user groups? PP1 and PP2 are related to how to explain and interpret PM outputs. Interpretability and explainability are beneficial to instill feelings of understanding within the system's users. There are numerous generic explainability and interpretability definitions [8] as it is becoming a popular research area. Especially in the field of artificial intelligence (AI) and in relation to machine learning, which use algorithms learned through their own training process. The 'black box' effect can lead to questions such as how the system made a particular decision. However, no such definition appeared to exist for general PM, e.g. process discovery, except for predictive process monitoring in recent research [9] or some focused use cases [10, 11].

RQ2: What are the main factors that can drive XPM and NPM within the certain phases of PM journey for certain user groups? There are many different drivers that can somehow affect both XPM and NPM. For example, educational aspects and PM experience of *individual* level during the initial PM phases (e.g. requirements gathering) or choices of visual components by the PM analyst during the technical development phases (e.g. dashboard building) of *technological* level.

RQ3: How can XPM and NPM be improved for certain PM stages and user groups for more efficient and insightful use of PM outputs? There are some existing methods and tools proposed by PM vendors or in the recent research [12] that aim (in-)directly to improve XPM and NPM. However, new solutions can be identified and existing ones can be (re-)evaluated. As mentioned in PP3, enhanced explainability and interpretability for more complex PM output(s) can contribute to better PM user adoption and value identification.

3. Research Methodology

We adopt a mixed methodology, combining qualitative and quantitative approaches. To address the research questions, we plan to use the four main phases shown in Figure 1.

To start with **definitions** (RQ1), we selected the initial scope. We will focus on three main *user groups* - business users, analysts, developers. We target planning, requirements and data collection, data processing and transformation, mining and analysis, evaluation, and process improvement and support as the most important high-level *PM phases* [13].

Once possible definitions per certain stages and certain user groups are clarified (RQ1), we plan to investigate **drivers** (RQ2) that can influence both XPM and NPM. For now, we have grouped these drivers into four categories. The categorization is preliminary and partially based on Technology-Organization-Environmental (TOE) framework [14] adding the individual perspective.

- **Personal / Individual** - level of education, personal background, level of experience in process mining, personal skills, etc.

- **Organisational** - change culture, learning culture, organizational structure and defined PM roles, etc. By organisations, we mean not only companies using PM software, but also others, e.g. PM suppliers and PM consulting firms.

- **Environmental** - industry and market structure, government regulation, etc.

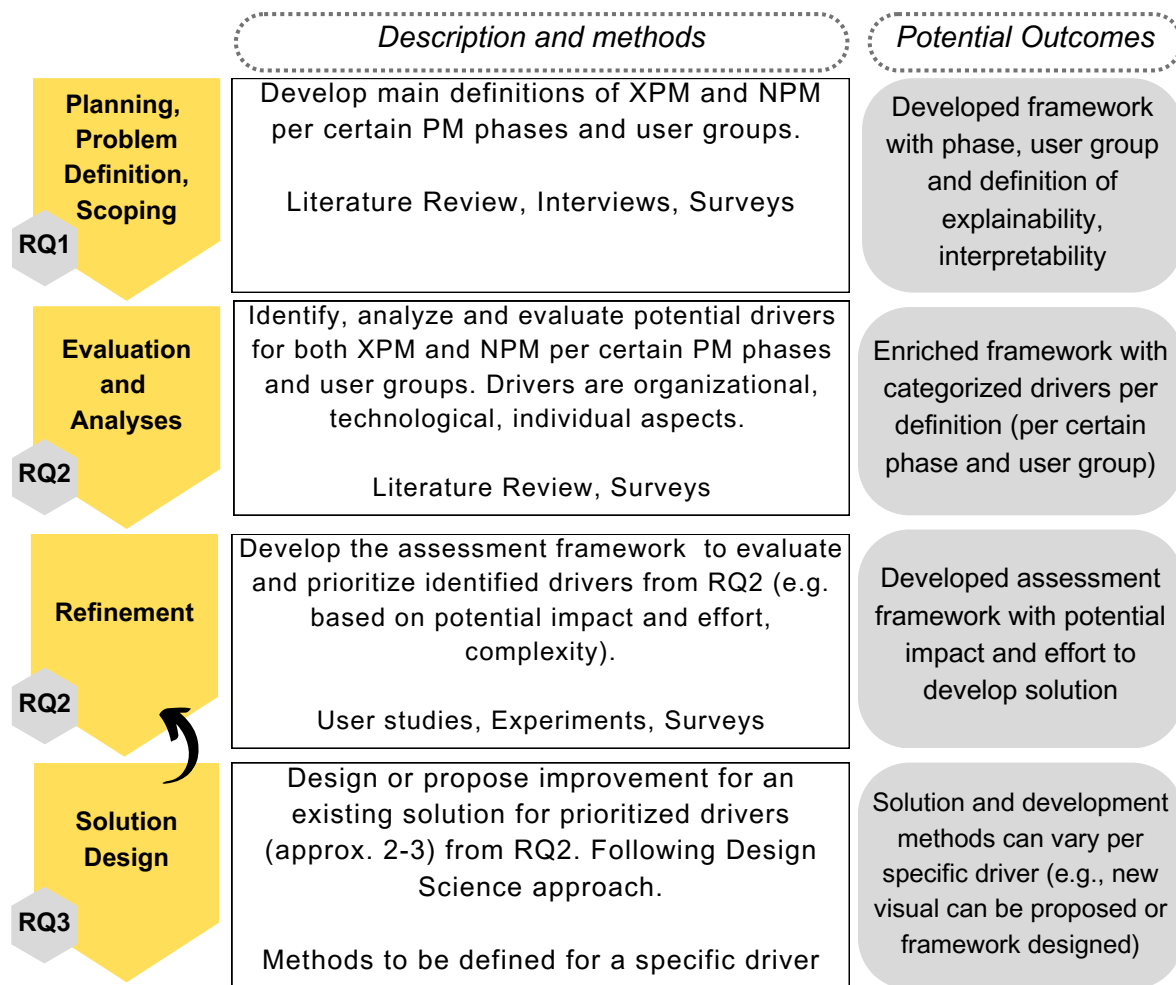


Figure 1: Initial Research Project Plan

- **Technological** - tooling and software (source systems, ETL solutions), visual components within those systems, programming languages and developed algorithms, etc.

After clarifying the definitions (RQ1) and potential drivers (RQ2), we aim to review, assess and prioritize them with the aim of identifying the drivers with the greatest business impact based on potential value. For two-three of the most impact drivers, we plan to review available **solutions** and evaluate them. Next, we intent to design improved solutions (RQ3) adopting a Design Science Research (DSR) approach [15]. For example, if the biggest impact (according to the chosen assessment) lies in the interpretability by the business user during the discovery phase (mining and analysis), a possible solution may be a new visual component or representation of PM output.

4. Current State

As preliminary research to answer RQ1, we conducted survey-based interviews. During the pilot interviews, we asked several process mining experts and 20 PM business users what do they understand by explainability and interpretability within process mining. When asking the questions ‘How would you define explainability in Process Mining?’ and ‘How would you define interpretability in Process Mining?’, we obtained the following results:

- Some respondents answered that they think explainability and interpretability are very important topics, but that they also find it difficult to answer one or both questions and propose a definition in a ‘constructive’ way.

- Others shared that explainability and interpretability are the same and that there is no difference in a definition, meaning they are interchangeable.
- In contrast, some emphasised the difference and even highlighted that one or the other is more important (e.g. that without explainability, there is no understandability and consequently no need for interpretability).

Initial results show that participants find XPM and NPM very important topic. However, some of them find it challenging to formulate clear definitions.

References

- [1] W. M. P. van der Aalst, *Process Mining - Data Science in Action*, Second Edition, Springer, 2016.
- [2] V. Mehta, M. Kerremans, C. Tornbohm, H. Chibber, Market share analysis: Process mining software, worldwide, Gartner Report (2023).
- [3] V. S. Dani, H. Leopold, J. M. E. M. van der Werf, I. Beerepoot, H. A. Reijers, From loss of interest to denial: A study on the terminators of process mining initiatives, in: CAiSE, volume 14663 of LNCS, Springer, 2024, pp. 371–386.
- [4] T. Grisold, J. Mendling, M. Otto, J. vom Brocke, Adoption, use and management of process mining in practice, *Bus. Process. Manag. J.* 27 (2021) 369–387.
- [5] F. Zerbato, P. Soffer, B. Weber, Process mining practices: Evidence from interviews, in: BPM, volume 13420 of LNCS, Springer, 2022, pp. 268–285.
- [6] V. S. Dani, H. Leopold, J. M. E. M. van der Werf, H. A. Reijers, Progressing from process mining insights to process improvement: Challenges and recommendations, in: EDOC, volume 14367 of LNCS, Springer, 2023, pp. 152–168.
- [7] G. Janssenswillen, B. Depaire, Towards confirmatory process discovery: Making assertions about the underlying system, *Bus. Inf. Syst. Eng.* 61 (2019) 713–728.
- [8] J. Droste, H. Deters, R. Fuchs, K. Schneider, Peeking outside the black-box: AI explainability requirements beyond interpretability, in: REFSQ Workshops, volume 3672 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024.
- [9] N. Mehdiyev, P. Fettke, Explainable Artificial Intelligence for Process Mining: A General Overview and Application of a Novel Local Explanation Approach for Predictive Process Monitoring, Springer International Publishing, 2021, p. 1–28. URL: http://dx.doi.org/10.1007/978-3-030-64949-4_1. doi:10.1007/978-3-030-64949-4_1.
- [10] T. Hoogendoorn, J. J. Arachchige, F. A. Bukhsh, Survey of explainability within process mining: A case study of BPI challenge 2020, in: FIT, IEEE, 2023, pp. 43–48.
- [11] L. Nannini, Explainability in process mining: A framework for improved decision-making, in: AIES, ACM, 2023, pp. 975–976.
- [12] S. Miksch, Visual analytics meets process mining: Challenges and opportunities, in: ICPM, IEEE, 2021, p. xiv.
- [13] M. L. van Eck, X. Lu, S. J. J. Leemans, W. M. P. van der Aalst, PM²: A process mining project methodology, in: CAiSE, volume 9097 of LNCS, Springer, 2015, pp. 297–313.
- [14] J. Baker, *The Technology–Organization–Environment Framework*, Springer New York, New York, NY, 2012, pp. 231–245. doi:10.1007/978-1-4419-6108-2_12.
- [15] P. Johannesson, E. Perjons, *An Introduction to Design Science*, Springer, 2014.