Survey of Explainability within Process Mining

Tjalling Hoogendoorn
University of Twente
Enschede, Netherlands

ABSTRACT

In 2017, Wil van der Aalst wrote a conference introduction in which he calls for more transparency and explainability within process mining. To understand this problem and the criticality of it, a literary analysis is done on the 37 papers submitted for the BPI 2020 challenge, observing the reported use of techniques and platforms. We gathered this information for all 37 papers in a table and found that only 3 papers report the techniques and settings used. This was a small scale survey on a very limited sample, further research is required in a more systematic way on a bigger sample size.

KEYWORDS

Data Science,Process Mining, Literature Review, Process Mining Transparency

ACM Reference Format:
Tjalling Hoogendoorn. 2023. Survey of Explainability within Process Mining. In *Proceedings of Twente Student Conference on IT (TScIT 38).* ACM, New York, NY, USA, 9 pages. https://doi.org/none

1 INTRODUCTION

Since computers have been able to automatically log and store data, data has been produced in ever greater quantity [12]. With the help of computers we have been able to navigate the flow of time to analyze events from completely new angles and visualize a previously unseen dimension ([1]P. 1-4). One of these techniques which has appeared in the last decade to analyze the flow of events we can now track is Process mining (pm). Pm is a technique to create a link between event logs and the process models that those logs came from ([12]P. 17). It is a method often applied in the analysis of business processes for the visual analysis of processes and to identify bottlenecks and anomalies within the process ([12] P. 25,[2][6]). However there are still many challenges when trying to visualize complex processes for analysis which often lead to complex models difficult to understand with the naked eye. However the easy interpretation of process deviations is an important aspect when trying to apply process mining in for example the fields of business and medicine[15][8]. As such there is a need for ”explainable” process models, which take complex data and models and provide a simple high level overview with the relevant data in a way that is easy to comprehend for humans.

TScIT 38, Feb 03, 2023, Enschede, Netherlands

ACM ISBN none.
https://doi.org/none

There is not 1 silver bullet when it comes to analyzing an event log. When analyzing temporal data, one can apply different levels of filtering, mapping and rendering to get different views of the same process ([1]P. 7). Different pm algorithms and their configurations will give different views of the same data [2]. 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) [2]. Especially during process discovery it is important to have a visual to uncover the patterns of the data that is being analyzed. However as processes can vary wildly, it is up to the analyst to find the best tool for the job, and new process mining techniques are appearing every year to provide new possibilities.

1.1 Background Information

To help with this research, data is used that is published as part of the IEEE process mining task force (tfpm) Business Process Intelligence (BPI) challenge [10]. The tfpm has as goal to promote the research into and adoption of process mining in the wider academic and professional community. As part of this goal the tfpm organizes the BPI Challenge where they invite anyone to mine valuable data insights from the provided datasets using whichever open source our proprietary pm algorithms they want and then provide an award for the best student and non-student papers at the Process Mining Conference. Using one of these datasets allows for a complex process in a single context in which pm analysis can be performed and different techniques can be tested.

Pm has so far been discussed in a very general focus, however there are many different focus points and angles from which one can apply pm research. Due to the nature of the data, there will be a limited focus on analysing data traces from an event log after a process has finished (considered a post-mortem), which stands in contrast to analysing a process while it is still happening (pre-mortem). Furthermore the focus of the analysis will be entirely on discovering the *de facto model*, or the model that describes the process as it happens in reality (process discovery) and possibly detecting bottlenecks and inefficiencies within the process (process enhancement), which stand opposed to anomaly detection by comparing a model of how the process should go with the model we get (Conformance checking) or giving advice or predictions on how to improve a process (Operational support, often closely tied with pre-mortem analysis). For more information, see part II.C from [2].

Explainability is a topic often discussed in the context of AI. The AI is a black box that makes decisions, however we cannot see how it came to the decision that it made. Explainability is the art of making the decision process transparent and giving context for a human to interpret the decision made by AI[5]. All of this is done in the hope of adding transparency to increase the acceptance of AI technology in society. However this is not a challenge that is faced by AI alone. The upsurge in advanced Machine Learning models, powered by large amounts of gathered and curated data, mean Data Science as a whole now faces these challenges of transparency and correct use of data. Pm also does not escape from this, as it is more and more used to predict and analyze processes and steer them to a certain way, we must be aware to make responsible use of the data we receive and use [13]. As such we also face a challenge within the field of pm to have transparent pm processes and data interpretation if we hope this field between process analysis and data science will remain accepted by the public as it becomes more known.

1.2 Problem Statement

PM as a new technology requires to be understood by non-practitioners to enable potential widespread adoption in the future[13]. To know what is and is not imported to be explained insight is needed in what methods/algorithms are used by practitioners, how much open and closed source platforms are used and how transparent reports are in disclosing used methods. We finally wish to answer how interpretability relates to explainability and how much explanation is necessary for a clear interpretation.

1.2.1 Research Questions. To find the answer to the problem statement, it is important to find out what the current state of explanation in PM reports is. specifically the following research questions are of interest:

- RQ1. What is the trend of process mining techniques/algorithms used?
- RQ2. What is the trend of platform used for process mining?
- RQ3. What is the prevalence of explaining process mining techniques?
- RQ4. What is the ratio of open source versus closed source platforms?

2 RESEARCH APPROACH

Our evaluation is based off off the PM^2 model described by van der Aalst and colleagues [14]. However since this research falls more in the category of a literature review rather than an application of process mining, The iterative approach taken was informed by methodology used in past literature reviews in Process mining by Garcia et al.(2019)[3], Thiede et al.(2018)[11] and Ghasemi and Amyot (2016) [4]. More rigorous approaches such as PRISMA [7], which has been developed for medical literature reviews and alternative approaches such as described by snyder(2019)[9], have not been chosen in favour of the simpler approach of past PM reviews, due to the small scope of this review. Since no literature methodology has been published for the domain of process mining, we end up with a Semi systematic general literature review approach based on the above methodology from Garcia et al however it cannot be called an official methodology.

2.1 Approach on Data Collection

For determining which techniques are commonly used within BPI[1], an evaluation was done of all 37 papers submitted. BPI challenge data has often been reused for research in the past[11]. Each paper is evaluated on the following conditions:

- Which platform is used?
- Which techniques are named by the paper as used to create the graphs used?

Then a final table will be made, focusing exclusively on papers which (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, both of which are closed source tools).

2.2 Approach on Data Processing

After all the data was gathered, they where categorized in which main platform was used and a frequency analysis on 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 regards to what was used for pm analysis:

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

The result of this can be observed in 1
Furthermore, The following unique categories are defined for frequency analysis:

- Celonis
- Disco
- ProM
- PM4Py
- Non-PM methods
- other (own software, other libraries)

The results of which can be observed in 2

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

3 RESULTS

A total of 37 papers have been submitted for the ICMP conference BPI challenge 2020. Of these papers, 3 papers used no modern process mining techniques, constructing BPM or Petri net graphs from raw statistical data instead (marked as using an unstated platform). 7 papers have a platform and results but no explanation on how the researchers got to these results (empty techniques field) and 11 papers have a platform but do not state the algorithms used, which are marked as simply using an Ideal Process Graph or throughput/replay analysis. Of the final 16 papers 13 simply mention or show using an "ideal process graph" or process visualisation/throughput analysis. table 1 shows the three papers which used (primarily) ProM and explicitly stated which techniques where 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 1: Three papers having a clear technique description and using open source software.

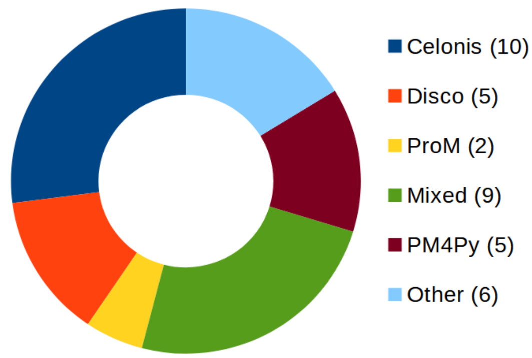

Figure 1: Donut chart showing which single platforms are prefered vs an equal mix of platforms of BPI 2020 papers

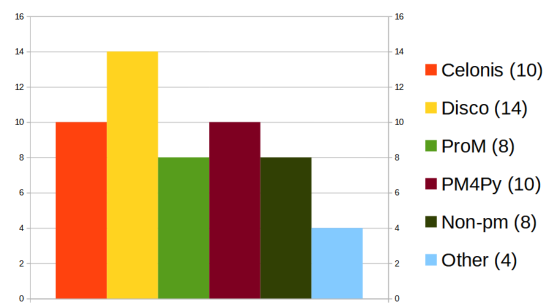

Figure 2: Frequency histogram of total use of platforms BPI 2020 papers

3.1 Observations on Platform Use

It can be observed that the 4 main Techniques used by prom From our dataset we can observe how often each platform is used and which platforms are most prevalent as a main platform in figures 2 and 1. The difference here is that certain papers use an equal mix of different platforms for different analysis (the total use of which is analyzed in figure 2). However some papers clearly favor one platform for the majority of their analysis, which is visualized in figure 1.

From figure 1 We can observe that Celonis is most often used as a primary or sole platform used in research. However we observe that a lot of Practitioners use a mix of two or more platforms in equal measure.

From figure 2 We can observe that Disco is the most frequently used platform, with Celonis and PM4py being second.

3.2 Result Analysis

From the above results we can attempt to answer our research questions. A 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.

From figure 1 we can see 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.

When we take into account our frequency analysis in figure 2 on We notice the most popular platform when including mixed use is actually Disco. Disco Visual analysis pops up in all mixed platforms (while ProM is able to provide the same function, Disco is favored here). It appears there is a trend to use Specifically Disco's Visual miner for early data exploration.

From the initial description. We can see that it is very, very rare for a paper to actually mention which specific miner has been used, as only 3 of the 37 papers do. 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 graph is made, 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, let us look at the ratio of open and closed source platforms. by again looking at figure 2. We find 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 our sample space. However when we observe figure 1 We do notice that open source platforms are rarely used as a primary platform for the bulk of analysis. Thus from here we can conclude that closed source platforms are slightly more popular, due to their ability to often be all you need or simply ease of use. This means that the usage of closed source platforms cannot be the only factor explaining why so few papers actually note how their process model is generated.
Finally when we wish to look at trends for specific process mining algorithms, since only 3 papers in the sample of 37 mentioned their techniques explicitely, we must conclude 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.

4 DISCUSSION

When gathering the data from the BPI challenge papers, a lot of papers show results but never mention their methodology. This means that a lot of papers that primarily used ProM where still disqualified on the second point, which is that they never stated which miners they used to get their graphs.
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 which can have influenced the study. What is clear is that there is a strong variety in the quality of the papers, and that despite a clear template being provided many submissions did not follow the submission template, with some having major deviations in style.
Due to lack of information, it is unclear if the lack of explanation a lot of papers displayed is also indicative of a lack of being able to understand these papers.
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.

4.1 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 and professionals who participated because their company mandates it mixed in with those that are actually interested in the BPI challenge intrinsically. 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 in the quality of PM reporting by practitioners in the field, but can be Interpreted as an indication.

4.2 Future Research

During the course of this research, the following avenues for potential followup research have been identified:

- Broadening the scope of papers covered in the review
- Making more fine grained meta observations on professional vs student entries, explaining data discovery, 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 followup integrative review to see how lessons from psychology and/or Explainable AI can be applied in this field.

5 CONCLUSION

This research paper set out to analyse explainability within process mining, to get better insights into where the focus should lay in 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 in the analyzed sample provided inconclusive data.
Important conclusions from this paper are that in this 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 simply 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 regards to open or closed source. There is however a trend within this sample to not 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.

6 ACKNOWLEDGEMENT

The author would like to thank Professor F.A. Bukhsh of the faculty of Datamanagement & Biometrics from the university of twente for their guidance and supervision during the research project. Furthermore the author would like to thank professor Y.J. Erden of the faculty of Philosophy at the university of twente for helping place the ethical discussion surrounding big data in context. Their insights where invaluable when considering the ethics that lay at the foundation of this paper.

REFERENCES

[1] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. 2011. *Visualization of Time-Oriented Data*. Springer London, London. https://doi.org/10.1007/978-0-85729-079-3

[2] Najwa Chaydy and Abdellah Madani. 2019. 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)*. IEEE, Casablanca, Morocco, 1–9. https://doi.org/10.1109/SysCoBIoTS48768.2019.9028024

[3] Cleiton dos Santos Garcia, Alex Meincheim, Elio Ribeiro Faria Junior, Marcelo Rosano Dallagassa, Denise Maria Vecino Sato, Deborah Ribeiro Carvalho, Eduardo Alves Portela Santos, and Edson Emilio Scalabrin. 2019. Process Mining Techniques and Applications – A Systematic Mapping Study. *Expert Systems with Applications* 133 (Nov. 2019), 260–295. https://doi.org/10.1016/j.eswa.2019.05.003

[4] Mahdi Ghasemi and Daniel Amyot. 2016. Process Mining in Healthcare: A Systematised Literature Review. *International Journal of Electronic Healthcare* 9, 1 (2016), 60. https://doi.org/10.1504/IJEH.2016.078745

[5] Leilani H. Gilpin, David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. 2018. Explaining Explanations: An Overview of Interpretability of Machine Learning. In *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, Turin, Italy, 80–89. https://doi.org/10.1109/DSAA.2018.00018

[6] Silvia Miksch. 2021. Visual Analytics Meets Process Mining: Challenges and Opportunities. In *2021 3rd International Conference on Process Mining (ICPM)*. IEEE, Eindhoven, Netherlands, xiv–xiv. https://doi.org/10.1109/ICPM53251.2021.9576854

[7] David Moher. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine* 151, 4 (Aug. 2009), 264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

[8] Morteza Noshad, Christian C. Rose, and Jonathan H. Chen. 2022. 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* 127 (March 2022), 104004. https://doi.org/10.1016/j.jbi.2022.104004

[9] Hannah Snyder. 2019. Literature Review as a Research Methodology: An Overview and Guidelines. *Journal of Business Research* 104 (Nov. 2019), 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

[10] TFPM. 2020. BPI Challenge – Process Mining Conference 2020.

[11] Malte Thiede, Daniel Fuerstenau, and Ana Paula Bezerra Barquet. 2018. How Is Process Mining Technology Used by Organizations? A Systematic Literature Review of Empirical Studies. *Business Process Management Journal* 24, 4 (Jan. 2018), 900–922. https://doi.org/10.1108/BPMJ-06-2017-0148

[12] Wil van der Aalst. 2016. *Process Mining*. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-49851-4

[13] Wil M. P. van der Aalst. 2017. Responsible Data Science: Using Event Data in a "People Friendly" Manner. In *Enterprise Information Systems*, Slimane Hammoudi, Leszek A. Maciaszek, Michele M. Missikoff, Olivier Camp, and José Cordeiro (Eds.). Vol. 291. Springer International Publishing, Cham, 3–28. https://doi.org/10.1007/978-3-319-62386-3_1

[14] Maikel L. van Eck, Xixi Lu, Sander J. J. Leemans, and Wil M. P. van der Aalst. 2015. PM$$2□$$: A Process Mining Project Methodology. In *Advanced Information Systems Engineering*, Jelena Zdravkovic, Marite Kirikova, and Paul Johannesson (Eds.). Vol. 9097. Springer International Publishing, Cham, 297–313. https://doi.org/10.1007/978-3-319-19069-3_19

[15] Anton Yeshchenko, Claudio Di Ciccio, Jan Mendling, and Artem Polyvyanyy. 2022. Visual Drift Detection for Event Sequence Data of Business Processes. *IEEE Transactions on Visualization and Computer Graphics* 28, 8 (Aug. 2022), 3050–3068. https://doi.org/10.1109/TVCG.2021.3050071

A FULL DATA TABLE

The following table has all 38 submissions from the ICPM 2020 BPI challenge,In the same order as the papers have been handed in (an earlier submission number means the paper has been handed in earlier).

ICPM 2020 sub-mis-sion	platform	Techniques Used	Other analysis done	Authors
44	ProM, Disco	ActiTraC model based clustering, Petri net Visualization, disco replay visualization	Compares Linear prediction and H2O autoML to predict spenditure from log attributes using R and Python	Stijn Kas, Ruben Post and Sebastiaan Wiewel
88	Celonis	ideal process, identifying and comparing common variations	histograms and mean value comparison.	Valentina Barrera, Sofia Redondo, Rosario Rodriguez, Juan Ignacio Silva, Ambrosio Valdes and Victor Galvez
89	Celonis	Conformance Checking, Ideal process, Visualisation with Celonis process explorer		Eleuterio Ramirez, Jose Luis Haddad, Maximiliano Stuardo, Juan Jose Martinez, Katherine Vergara, Benjamin Rivera and Victor Galvez
90	Celonis	Visualisation with Celonis Process Explorer, ideal process and conformance checking	Marker charts of throughput, histogram of rejection flows	Nicolas Acosta, Roshad Alipanah, Vicente Etchegaray, Francisca Ibarra, Flavio Tarsetti and Victor Galvez
98	Disco	Ideal process graph, basic model, conformance checking using outliers identified with boxplots	Histogram with correlation between actors and rejection. Average spending over or under budget.	Wessel van Bakel, Rose Mary Hulscher, Mitchell Klijs and Martijn Sturm
99	Python RMM4Py and PM4Py libraries			Sabine Klein, Johannes Lahann, Lea Mayer, Dominic Neu, Peter Pfeiffer, Adrian Rebmann, Martin Scheid, Brian Willems and Peter Fettke
100	Disco	Ideal process graph, Main deviation		Alexander Nikolayuk, Anna Shtokolova, Evgeniya Sdvizhkova, Yulia Khabarova and Yuri Tarasov
105	Unstated	Frequency graph	DBSCAN cluster to identify most popular process types., making bpmn graphs by hand	Igor Kashirin, Vladimir Sudarev, Elena Popova, Dmitry Usachev, Makar Shadiyan, Anastasia Knishenko and Daniil Surnyaev
106	Celonis	Ideal process graph	Histograms and boxplots in python, uses sql queries to answer most questions	Stanislav Belov, Aleksandr Shevchenko, Roman Mushkudiani, Nataliya Mityunina, Anna Levkovskaya and Evgenii Lebedev
108	pm4py	T-SNE		Dmitrii Khodaev, Viktor Kalugin, Evgenia Korneeva, Marina Savintseva and Anastasiya Balashova

Table 2

ICPM 2020 submission	platform	Techniques Used	Other analysis done	Authors
111	Disco	ideal process, mean throughput, Process Visualisation		Igor Zhilkin, Maksim Deyneko, Margarita Fadyushina, Roman Iov, Vladislav Makarov, George Alyaev, Vitaly Bordachev, Natalia Karpova, Vadim Nikolaev, Igor Nizamov, Konstantin Shusterzon, Anton Stanevich and Yuri Tiguntsev
120	myPM(self-dev), sql	sql queries		Dmitry Trifonov, Khabarova Ekaterina, Kalinin Yuriy, Evtushenko Aleksandr and Zhemchuzhnikov Evgeny
121	Celonis			Anna Abrashkina, Anastasiya Golovataya, Andrew Farhutdinov and Alexey Zuravlev
122	SQL, own software			Betale Pasliokob
124	Disco, Einstein Analytics, pm4py	T-SNE, DBSCAN, Kmeans		Sergey Tsaplin, Ekaterina Skvortsova, Anastasiya Paklieva, Nikita Zakoryuchkin, Danila Kostin, Vsevolod Zarubin, Oleg Neelov, Dmitrii Bityutskikh, Olga Fokina, Aleksandr Veselov, Aleksandr Shevchenko, Aleksandr Isaev, Aleksandr Kurilenko, Denis Korneev, Mihail Kargin and Maksim Kataev
125	Python, SQL, Celonis	ML to predict case length		Stanislav Laptenkov, Konstantin Grishchenko, Sergey Zakharov, Vyacheslav Chernov, Artem Kropis, Roman Kozhushko and Sergey Lebedev
126	ProM	Visual Analysis, Inductive, Heuristic and Fuzzy Miner, Bonnet Test, Pareto analysis		Antonio Davide Ciappina
127	Disco, Python		Uses python statistical libraries and Disco Ideal process Graphs and common deviations.	Mariya Devyatova, Svetlana Zverintseva, Tatyana Senicheva, Elena Puchnina, Ekaterina Danilovich, Marina Ivanova, Svetlana Stroganova, Panova Natalya and Kseniya Golovina
128	Celonis, Statistica			Alyona Stepanova, Evgeny Igumnov, Maxim Milovanov, Dmitry Samus, Andrey Slutskiy, Nikolay Nikolaev and Ilya Mossur

Table 3

ICPM 2020 submission	platform	Techniques Used	Other analysis done	Authors
129	ProM, PM4Py, Excel	Excel for exploration, Directly follow graphs	Unclear exactly what algorithms are used as it is not clearly stated in the paper.	Alexey Myasoedov, Galina Vlasova, Tatyana Seregina and Sofia Ionova
130	PM4Py	Directly-follow graphs	Welch’s t-test, box plot visualization, histograms	Mikhail Poruchikov and Pavel Katkov
131	PM4Py (unclear)	Heuristic miner (dependency thresh 0.999)		Georgy Buzikashvili
132	Disco, ProM, PM4Py	(disco) throughput and play-in map, other Unknown	One-class Support Vector Machine (Sci-kit)	Elena Samtsova, Anna Vishnivetskaya, Murtazali Murtazaliev, Andrey Starchenkov, Anna Martynova, Alena Surzhikova, Victoria Kapishina and Elena Alymova
134	Celonis	Throughput analysis, ideal process flow		Olga Sidorkina and Maria Kosyrkova
136	Aris express, Disco	Process Visualization, sql queries.		Ruslan Filipov, Roman Krekhno, Anastasiya Sviridenko, Alexander Balandin and Evgeniya Shtina
137	Disco			Anton Zheronkin, Ivan Ogurtsov, Pavel Kalashnikov, Dmitry Lukyanenko, Sergey Gevorkyan and Nikita Altuhov
141	sberpm python,	Autoinsights algorithm, Heuristic miner		Aydar Bulatov, Anna Sverkunova, Artem Glagolev, Elvares Geydarov, Sergey Kuznetsov, Elena Khomyakova and Danil Smetanev
142	celonis			Luis Armando González, Hernan Arenas, José Lara, Francisca Rebolledo and Klaus Ribbeck
146	Disco, ProM, PM4Py	Inductive Visual Miner (ProM), Performance Spectrum Miner (ProM)		Dorina Bano, Maximilian Völker, Simon Remy, Henrik Leopold and Mathias Weske
147	PM4Py	Unstated		Paul Bogurenko, Andrew Stukolov and Alexander Teptyarev
148	Disco, Celonis, ProM	Ideal process graph, Conformance checking, Disco process replay		Luise Pufahl, Richard Hobeck, Paul Binetruy, Wepan Chada, Mykola Digtiar, Kim Julian Gülle, Marta Slarzynska, Fabian Stiehle and Ingo Weber

Table 4

ICPM 2020 submission	platform	Techniques Used	Other analysis done	Authors
150	ProM	Inductive Visual Miner (ProM), Directly follow graph(ProM)		Georgiy Zakharov, Aleksanrda Lipunova, Aleksey Makhov, Alexandr Shpitalnik and Oleg Pavlyukov
152	Disco, Prom	Top 5 process variations	Boxplots, scatterplots	Kushal Poddar, Monika Gupta, Jay Bandlamudi and Sampath Dechu
156	Disco	Ideal process graph, Common deviations, replay analysis		Tatyana Zolotova, Konstantin Mikiev, Daria Zaplotnikova and Kseniia Khandogina
157	unstated	unstated		Aleksandra Piasecka, Paul Giessler and Oskar Leligdowicz
177	unstated	BPMN Models	Comparison of attributes in R, correlation maps, made bpmn graphs by hand	Chiao-Yun Li and Jingjing Xu
179	Disco, PM4Py			Victor Chufistov, Yuliya Trushik, Nikita Uliashenkov and Sergey Cherenkov

Table 5