M.J. Hornos et al. (Eds.)

doi:10.3233/AISE240009

On the Usefulness of Human Behaviour Process Models: A User Study

Silvestro VENERUSO [a,1], Francesco LEOTTA [a], and Massimo MECELLA [a]
[a] *Department of Computer, Control and Management Engineering*
Sapienza Università di Roma - Rome, Italy

Abstract. Last year witnessed a growing interest from the Business Process Management research community in analyzing activities carried out in sensorized environments using techniques originally intended for business processes. However, activities conducted in such scenarios differ significantly from typical processes in terms of repetitiveness and predictability. This raises the issue of assessing the suitability of state-of-the-art modeling formalisms and mining techniques to represent them, especially when humans are involved. In this paper, we present the results of a user study conducted with this specific goal. Specifically, we analyze the opinions of a group of experts regarding different representation formalisms and mining algorithms, drawing conclusions about the usefulness of such models in smart environments.

Keywords. smart homes, process discovery, habit mining, user study

1. Introduction

In recent years, there has been a notable surge in interest among researchers and manufacturers regarding the application of Process Mining (PM), particularly focusing on process discovery techniques, to analyze human behavior within smart spaces. Throughout the remainder of this paper, we use PM and process discovery interchangeably as synonymous terms. The wealth of data automatically gathered through IoT sensors serves the purpose of gaining insights into user behavior, such as sleep tracking, or executing automated actions on behalf of the user, such as automatically adjusting blinds. Noteworthy examples of current applications for monitoring human behavior in smart spaces include the utilization of smart thermostats like the Google Nest Learning Thermostat and ambient assisted living systems, such as those designed for detecting falls in the elderly.

Both Process Mining (PM) and smart spaces have undergone rapid evolution as distinct fields of study. However, researchers have started to explore the synergies between these disciplines, yielding intriguing results that warrant thorough analysis and comparison for future advancements. The application of PM techniques to data from smart spaces facilitates the modeling and visualization of human habits as processes. Nevertheless, despite the potential extraction of process models from smart spaces data, several significant challenges have surfaced when adapting techniques originally designed for business

[1] Corresponding author: Silvestro Veneruso, veneruso@diag.uniroma1.it

processes (BP) to the study of human behavior [1]. These challenges include: (1) selecting or designing an appropriate modeling formalism for representing human behavior, (2) bridging the gap between sensors and event logs, (3) segmenting logs into traces to facilitate the application of PM techniques, (4) addressing the complexities of multi-user environments, and (5) tackling the continuous evolution of human behavior.

Process mining [2] provides several modeling approaches, each offering unique insights into the representation of underlying processes. In particular, *process discovery* is a process mining technique used to discover and generate the process model describing the underlying behavior shown in the log.

Starting from the results obtained by applying two state-of-the-art unsupervised methodologies that segment the log on a habit basis [3] and on an activity basis [4], we are interested in applying and comparing the output of three different *discovery* algorithms: the inductive, heuristic, and fuzzy miner.

The objective of this work is *(i)* to collect user feedback via questionnaire to establish which of these three discovery algorithms is most suitable for modeling human behavior in smart spaces and, consequently, *(ii)* to extend the validation of the approaches published in [3] and [4].

The paper is organized as follows: Section 2 introduces background concepts and terminology; Section 3 describes the procedure followed to conduct the user study; then, results are analyzed and discussed in Section 4; Section 5 introduces relevant related works; finally, Section 6 concludes the paper.

2. Background

In this work, we focus on the challenge of modeling human behavior in a smart environment. In particular, referring to the terminology described in [5], a model in a smart home may refer to different concepts:

- *action*, i.e., atomic interaction with the environment or a part of it (e.g., turning on the TV);
- *activity*, i.e., a group of human atomic interactions with the environment (actions) that are performed with a final goal (e.g., cleaning the house);
- *habit*, i.e., a group of actions or activities (one in the extreme case) that define what happens in specific contextual conditions (e.g., what the user usually does in the morning between 08:00 and 10:00).

Models of human habits and activities can be either manually defined (i.e., *specification-based*) or obtained through automated techniques (i.e., *learning-based*).

In specification-based methodologies, models are usually based on logic formalisms, which are relatively easy to read and validate (once the formalism is known to the reader), but their creation requires a major cost in terms of expert time and effort.

In the learning-based case, the model is automatically learned from a training set (whose labeling cost may vary according to the proposed solution), but employed formalisms are usually not "explainable" due to the statistical techniques they are based on, making them less immediate to understand [6]. The practical applicability of techniques proposed in the literature is limited by the effort required by the final user to manually label smart space logs. Approaches based on supervised (or weakly supervised) learning require the logs to be labeled with markers denoting the onset and end of all (or at least of a consistent subset) of the occurrences. However, manual labeling of logs is perceived by the final users as annoying, which could result in imprecise labeling, possibly tampering with the performance of algorithms at runtime.

Human-readable formalisms should be used alongside unsupervised machine learning techniques; this is the most important challenge in this field of research. Applying process mining to smart spaces allows you to get the best of both worlds because processes are human-readable, formally grounded, and can be mined automatically [1].

Process mining (PM) [2] is a fairly recent research discipline that combines data mining techniques with techniques used in business process management (BPM) [7]. Its main objective is to extract meaningful information from event logs. Among the several PM techniques, we are interested in *process discovery*, i.e., a technique for discovering the process model describing the behavior shown in the event log. Such a technique produces as output a process model most commonly represented using the Petri net formalism, i.e., a directed graph composed of nodes and arcs (respectively called *places* and *transitions*).

In [3], authors propose an unsupervised methodology allowing, given a sensor log, to automatically segment *human habits* by applying a classical bottom-up discretization strategy on the timestamp attribute. Such a class of discretization algorithms finds the best division of a continuous attribute by iteratively merging contiguous sub-ranges (also called “bins”) following a quality evaluation heuristic. In their proposal, the heuristic is based on quality measures computed on the process models automatically mined, through process discovery, from the intermediate bins. In particular, they drive the discretization targeting process models with high *simplicity* and low *structuredness*. Each obtained bin then represents a time range in which the human is supposed to perform activities following a clearly identifiable human process.

Similarly, in [4], authors introduce a fully automated log segmentation technique able to mark the beginning and end of each activity repetition in a sensor log. In order to obtain this result, the proposed technique employs the information about human position in the log to extract high-level actions (e.g., standing still or operating in a specific area of the house). Then, inactivity periods are analyzed in order to perform the first segmentation. Finally, clustering is employed to identify classes of segments representing activities. The work introduced in [3] only focuses on temporal-based segmentation targeted at defining *habits*. Conversely, this last work [4] focuses on activities instead of habits, which allows for finer-grained control over human routines.

3. Study Design

In order to mine the habit and activity process models from the segmentation results respectively obtained by applying the approaches described in [3] and [4], we have used the ProM software tool. ProM is an extensible framework that supports a wide variety of process mining techniques in the form of plugins (see: `https://promtools.org/`). It provides the latest implementations of the known process mining algorithms, not only for the *discovery* of process models but also for checking the conformance of a model and/or for enhancing it [8].
For each habit interval identified by the segmentation methodology described in [3], the ProM framework was used to discover the related process models. In particular, four

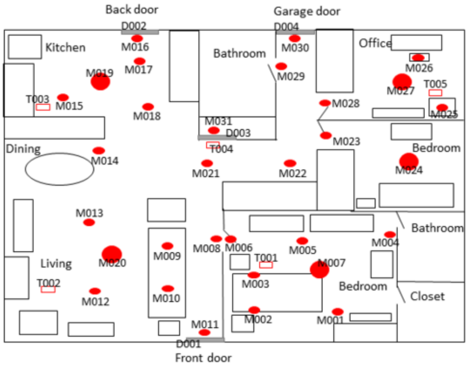


Figure 1. Aruba installation from CASAS project.

algorithms were used: *(i)* the inductive miner [9], *(ii)* the heuristic miner [10], *(iii)* the fuzzy miner [11], and *(iv)* the alpha miner [2] (this latter as a baseline). The models discovered with the alpha miner were excluded from the questionnaires because obtained results were below an acceptable quality.

Similarly, for each activity identified by the segmentation methodology described in [4], the ProM framework was used to discover the related process models. The same four discovery algorithms were used. The models discovered with the alpha miner were excluded from the analysis because they were not relevant.

We have applied both approaches to the Aruba dataset from the CASAS project[2]. It consists of a sensor log containing raw sensor measurements collected in a smart home inhabited by an adult woman for 220 days. The floorplan of this installation is shown in Figure 1 where available sensors are shown. In particular, the environment contains *(i)* Presence InfraRed (PIR) Sensors represented with small and large red ellipses and denoted with a label of the form MXXX, *(ii)* temperature sensors represented with red empty rectangles and denoted with a label of the form TXXX, and *(iii)* door switch sensors represented with gray rectangles and denoted with a label of the form DXXX.

In Figure 2, we propose some of the most relevant models mined from the ProM tool. All the other models are available at the link in the footnote[3].

Procedure. The user study was conducted following a questionnaire-based approach. The questionnaire was designed, created, and distributed to participants using the Google Forms platform.

Participants. Overall, a total of 20 different participants were involved in the user study. The age range was (on average) between 18 and 34 years and involved people with bachelor's degrees up to post-doc.

No previous knowledge was required to complete the questionnaire. However, before starting, it was recommended to read a brief handbook, which introduces the basic concepts and terminology for tackling the questionnaire. The handbook is available at the link in the footnote[4].

[2] see `http://casas.wsu.edu/datasets/`

[3] see: `https://drive.google.com/drive/folders/1-nj4jSWd3I1LC1vnT1cYqKh5Yhnxj-vE?usp=sharing`

[4] see: `https://drive.google.com/file/d/1szyCCN_bM_OGo2iTI4IPQPOInbeL52-B/view`

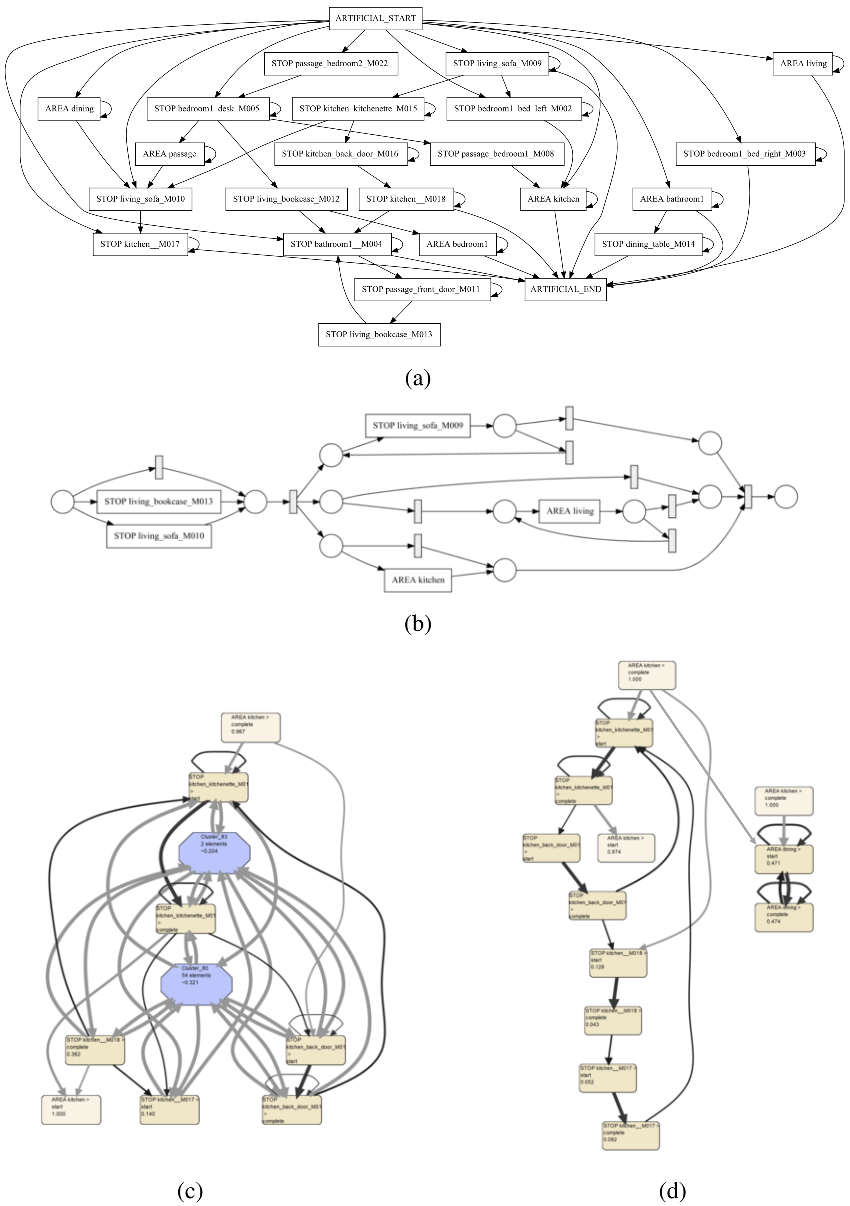


Figure 2. The figure shows: (a) the process model representing the "bed to toilet" activity extracted from the *heuristic* miner; (b) petri net of the "05:15-07:00" habit filtered on the activity "relax" extracted from the *inductive* miner; (c) a filtered component that emphasizes significant nodes in the model provided in the "wash dishes" model extracted from the *fuzzy* miner; (d) a filtered component that emphasizes significant nodes in the model provided in the "eating" model extracted from the *fuzzy* miner.

Questionnaire design. The questionnaire included 38 sections organized as follows:

- Section 1 contains a recommendation for reading the handbook before starting. The link to the handbook was provided.

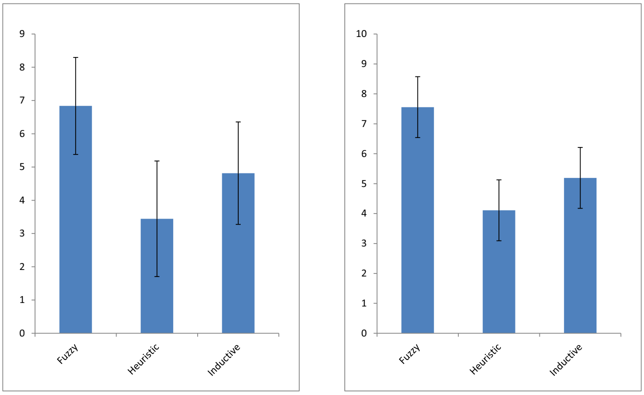
- Section 2 were about user profiling, i.e., age and current position.
- Sections 3 to 20 are designed to provide feedback on the process models mined over the activity-based segmentation results from the approach described in [4]. In particular, for each relevant activity, the related process model was mined by using three different *discovery* algorithms, i.e., *(i)* the inductive miner, *(ii)* the heuristic miner, and *(iii)* the fuzzy miner. Each output has been evaluated by visually inspecting the specific process model and by answering three questions:

  * **Question 1**: "*How well do you think this model reflects the activity x?*", where *x* was the activity under analysis in that specific section. It was rated on a Likert scale ranging from 1 ("too generic") to 10 ("too specific"). Here, we wanted to have a high-level feedback on the model in its entirety.
  * **Question 2**: "*Do the single actions in the model make sense with respect to a possible activity x?*", where *x* was the activity under analysis in that specific section. It was rated on a Likert scale ranging from 1 ("not at all") to 10 ("completely suitable"). Here, with respect to the previous question, we wanted to have a low-level feedback on the individual nodes (i.e., actions) of the model.
  * **Question 3**: "*Do you have any comments about this model?*". It was an optional, open-ended question to collect further feedback on the model under observation.

- Sections 21 to 38 are designed to provide feedback on the process models mined over the habit-based segmentation results from the approach described in [3]. In particular, for each habit, the related process model was mined by using three different *discovery* algorithms, i.e., *(i)* the inductive miner, *(ii)* the heuristic miner, and *(iii)* the fuzzy miner. Each output has been evaluated by visually inspecting the specific process model and by answering three questions:

  * **Question 1**: "*How well do you think this model reflects a possible daily human routine covering the time between START_TIME to END_TIME?*", where the time range was related to the habit under analysis in that specific section (e.g., from 05:15 AM to 7:00 AM). It was rated on a Likert scale ranging from 1 ("too generic") to 10 ("too specific"). Here, we wanted to have a high-level feedback on the model in its entirety.
  * **Question 2**: "*Do the single actions in the model make sense with respect to a possible daily human routine covering the time between START_TIME to END_TIME?*", where the time range was related to the habit under analysis in that specific section (e.g., from 05:15 AM to 7:00 AM). It was rated on a Likert scale ranging from 1 ("not at all") to 10 ("completely suitable"). Here, with respect to the previous question, we wanted to have a low-level feedback on the individual nodes (i.e., actions) of the model.
  * **Question 3**: "*Do you have any comments about this model?*". It was an optional, open-ended question to collect further feedback on the model under observation.

Questionnaire results. The feedback was collected in an Excel file and then analyzed with ad hoc statistical tools.



(a) (b)

Figure 3. Bar charts showing the mean scores of the responses to the various algorithms under analysis. Figure (a) shows the scores for question 1, while Figure (b) shows the scores for question 2.

4. Results and discussion

As described in Section 3, for each process model, two questions were asked: the first relating to the global model (high-level analysis) and the second specific to the nodes (i.e., human actions) included in the model (low-level analysis). Then, we ran two separate ANOVA tests, one for each question. The results are respectively shown in Table 1 and Table 2. In particular:

- Table 1 shows that there is a significant difference at the p-level for the three miners [F(1417.143 ; 2,426) = 198.90, p = 9.335E-69].
- Table 2 shows that there is a significant difference at the p-level for the three miners [F(1329.774 ; 2,128) = 212.82, p = 1.58914E-72].

These results revealed that there is a 100% chance that at least one discovery algorithm has a significant difference in mean scores.

In addition, post-hoc tests have been conducted to explore pairwise differences between the three discovery algorithms under observation, i.e., fuzzy, heuristic, and inductive. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the fuzzy condition (M = 7.557, SD = 1.791) was significantly different than the heuristic condition (M = 4.109, SD = 2.144) and the inductive condition (M = 5.192, SD = 1.696).

Taken together, these results suggest that process models mined using the *fuzzy* algorithm are considered more suitable for this type of modeling of human behavior, both at a high-level (i.e., question 1) and at a low-level (i.e., question 2).

The mean scores obtained from the questionnaire, shown in the bar charts in Figure 3, further highlight the participants' preferences.

Human behavior is flexible by nature, and this characteristic is considered by fuzzy mining. In particular, inductive process mining excels in capturing implicit knowledge, heuristic approaches leverage predefined rules and domain knowledge, while fuzzy mining accommodates uncertainty and imprecision in the data [12].

Table 1. Table (a) shows the results obtained by performing the 1-way ANOVA test on the feedback relating to the activity or habit model in its entirety (i.e., question 1). Table (b) shows the relevant calculations made to calculate these results.

Source	df	SS	MS	F	*p*-value
Factor (Between Groups)	2	1417.143	708.571	198.90	9.335E-69
Error (Within Groups)	681	2,426	4		
Total	683	3843.169			
F critical Value = 3.008949291					

(a)

	Fuzzy	Heuristic	Inductive	Total
Mean	7.557	4.109	5.192	5.619
Standard Deviation	1.791	2.144	1.696	2.372
Variance	3.208	4.600	2.878	5.626
t-critical	1.970	1.970	1.970	
Margin	1.283	2.084	1.467	
Grand Mean	5.619			
SS Total	3,843.2			
Sum of Squares Factor	855.567	520.023	41.551	
SS Factor	1417.143			
SS Error	2,426.03			

(b)

For the sake of brevity, individual comments have not been included in this papers. Anyway, participants highlighted limitations of both the evaluated approach at capturing some of the aspects of activities and habits.

5. Related works

The different formalisms, not only from the BPM area, employed to model human habits and activities in smart spaces are the subject of [5].

The possible applications of BPM and process mining in the world of IoT, thus also including smart spaces, are discussed in [13]. This is a manifesto, authored by a consistent part of the BPM research community, where involved challenges are described, and is the result of years of investigation (e.g., [1,14]) in specific areas of applications of IoT.

The application of process mining, and in particular process discovery, to the smart space scenario is the subject of [15]. In this survey, though, no user evaluation is conducted on the suitability of process modeling formalisms or discovery algorithms to be employed for visual inspection.

Table 2. Table (a) shows the results obtained by performing the 1-way ANOVA test on the feedback related to the specific actions included in the activity or habit model under observation (i.e., question 2). Table (b) shows the relevant calculations made to calculate these results.

Source	df	SS	MS	F	*p*-value
Factor (Between Groups)	2	1329.774	664.887	212.82	1.58914E-72
Error (Within Groups)	681	2,128	3		
Total	683	3457.292			
F critical Value = 3.008949291					

(a)

	Fuzzy	Heuristic	Inductive	Total
Mean	7.557	4.109	5.192	5.619
Standard Deviation	1.791	2.144	1.696	2.372
Variance	3.208	4.600	2.878	5.626
t-critical	1.970	1.970	1.970	
Margin	1.283	2.084	1.467	
Grand Mean	5.619			
SS Total	3,843.2			
Sum of Squares Factor	855.567	520.023	41.551	
SS Factor	1417.143			
SS Error	2,426.03			

(b)

The suitability of the different modeling formalisms to model human behavior in smart spaces is instead the topic of [16]. However, differently from this paper, the authors focus on a subjective analysis of the characteristics of the single languages, whereas a user evaluation is not provided.

6. Conclusions

In this article, we enriched the quantitative results already obtained from the unsupervised segmentation methodologies described in [3] and [4] with a qualitative analysis following a questionnaire-based approach. Subsequently, user feedback was analyzed with an ad hoc statistical tool, namely the analysis of variance (ANOVA) technique, in combination with further pairwise tests. The results show that among the three discovery algorithms used, the fuzzy miner is considered the most suitable for modeling human behavior (i.e., activities and habits), and this result is statistically significant.

The result obtained by the analysis raises important considerations related to the employment of human models extracted by using process mining, and especially process discovery, from IoT data. Fuzzy mining is a discovery technique (with an associated representation formalism) that is specifically intended for process performance analysis and visual inspection. It borrows concepts from cartography, allowing the final user to zoom in and out of a process, highlighting certain aspects of interest. This is certainly valuable for the smart space community, especially when applied to the case of the elderly, for which we want to analyze the daily routine.

One of the goals of smart space is also to automate, as much as possible and constrained to certain safety rules, human routines. From this point of view, fuzzy mining is less suitable as a representation formalism with respect, for example, to Petri Nets, which are instead discovered by the inductive and heuristic miners to model precise instructions. This is, on the one hand, an advantage, as fuzzy mining was originally intended to be effective for those processes that are flexible in nature, which include human processes. On the other hand, the kinds of suggestions that can be obtained from a fuzzy model are likely to be confirmed by the human inhabitant before enactment.

In general, the effective employment of human process models for enactment in smart spaces is an open research challenge from the point of view of Human-Computer (Home) Interaction, especially in light of the growing employment of home assistants available on the market.

References

[1] Leotta F, Mecella M, Mendling J. Applying process mining to smart spaces: Perspectives and research challenges. In: CAiSE 2015 Workshops. Springer; 2015. p. 298-304.

[2] van der Aalst WMP. Process Mining: Data Science in Action. 2nd ed. Springer; 2016.

[3] Esposito L, Leotta F, Mecella M, Veneruso S. Unsupervised segmentation of smart home logs for human habit discovery. In: 18th International Conference on Intelligent Environments (IE). IEEE; 2022. p. 1-8.

[4] Leotta F, Mecella M, Veneruso S. Unsupervised Segmentation of Smart Home Position Logs for Human Activity Analysis. In: 19th International Conference on Intelligent Environments (IE); 2023. p. 1-4.

[5] Leotta F, Mecella M, Sora D, Catarci T. Surveying human habit modeling and mining techniques in smart spaces. Future Internet. 2019;11(1):23.

[6] Hagras H. Toward human-understandable, explainable AI. Computer. 2018;51(9):28-36.

[7] Dumas M, La Rosa M, Mendling J, Reijers HA, et al. Fundamentals of business process management. vol. 1. Springer; 2013.

[8] Van Der Aalst W. Process mining: Overview and opportunities. ACM Transactions on Management Information Systems (TMIS). 2012;3(2):1-17.

[9] Leemans SJ, Fahland D, Van Der Aalst WM. Discovering block-structured process models from event logs containing infrequent behaviour. In: BPM'13. Springer; 2013. p. 66-78.

[10] Weijters A, Ribeiro JTS. Flexible heuristics miner (FHM). In: 2011 IEEE symposium on computational intelligence and data mining (CIDM). IEEE; 2011. p. 310-7.

[11] Günther CW, Van Der Aalst WM. Fuzzy mining–adaptive process simplification based on multi-perspective metrics. In: BPM; 2007. p. 328-43.

[12] Günther CW. Process mining in flexible environments. 2009.

[13] Janiesch C, Koschmider A, Mecella M, Weber B, Burattin A, Di Ciccio C, et al. The internet of things meets business process management: a manifesto. IEEE Systems, Man, and Cybernetics Magazine. 2020;6(4):34-44.

[14] Baier T, Mendling J, Weske M. Bridging abstraction layers in process mining. Information Systems. 2014;46:123-39.

[15] Bertrand Y, Van den Abbeele B, Veneruso S, Leotta F, Mecella M, Serral E. A survey on the application of process discovery techniques to smart spaces data. EAAI. 2023;126:106748.

[16] Di Federico G, Burattin A, Montali M. Human Behavior as a Process Model: Which Language to Use? In: ITBPM@ BPM; 2021. p. 18-25.