Bot Log Mining: Using Logs from Robotic Process Automation for Process Mining

Andreas Egger*[1][0000−0001−5580−4274], Arthur H.M. ter Hofstede[2][0000−0002−2730−0201], Wolfgang Kratsch[1][0000−0001−9815−0653], Sander J.J. Leemans[2][0000−0002−5201−7125], Maximilian Röglinger[1][0000−0003−4743−4511], and Moe T. Wynn[2][0000−0002−7205−8821]

[1] FIM Research Center, University of Augsburg and University of Bayreuth, Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Germany
{andreas.egger, wolfgang.kratsch, maximilian.roeglinger}@fim-rc.de
[2] Queensland University of Technology, Brisbane, Australia
{a.terhofstede, s.leemans, m.wynn}@qut.edu.au

Abstract. Robotic Process Automation (RPA) is an emerging technology for automating tasks using bots that can mimic human actions on computer systems. Most existing research focuses on the earlier phases of RPA implementations, e.g. the discovery of tasks that are suitable for automation. To detect exceptions and explore opportunities for bot and process redesign, historical data from RPA-enabled processes in the form of bot logs or process logs can be utilized. However, the isolated use of bot logs or process logs provides only limited insights and not a good understanding of an overall process. Therefore, we develop an approach that merges bot logs with process logs for process mining. A merged log enables an integrated view on the role and effects of bots in an RPA-enabled process. We first develop an integrated data model describing the structure and relation of bots and business processes. We then specify and instantiate a 'bot log parser' translating bot logs of three leading RPA vendors into the XES format. Further, we develop the 'log merger' functionality that merges bot logs with logs of the underlying business processes. We further introduce process mining measures allowing the analysis of a merged log.

Keywords: Robotic Process Automation · Process Mining · Business Process Management.

1 Introduction

Robotic Process Automation (RPA) is an emerging technology that refers to tools that mimic human actions on computer systems by interacting with the user interface or by connecting to APIs [3,28]. Applied for repetitive tasks, RPA can replace or even outperform humans regarding time, costs and quality [14,29]. It can be seen as the administrative counterpart of manufacturing robots [16].

* Corresponding author.

For example, with the help of RPA, Telefónica O2 improved 15 core processes in 2015 and achieved a three-year return on investment of 650%-800% [17]. Further benefits of RPA, like increased productivity, consistency and reliability, have been shown in literature [17,24,26,29]. RPA technology is already used by many organizations and the usage is expected to rise [7,18]

In the relatively new research field around RPA, some approaches combine techniques from process mining with RPA [20]. Process mining uses event logs recorded in information systems to discover, monitor, and improve business processes [2]. In the literature combining RPA and process mining, methods are proposed that discover the business processes that are best suited for automation [20,22]. Other concepts support the development phase of bots, e.g. by recording user actions to derive process models [21,23]. However, the post-implementation phase, i.e. when bots are already deployed in an organization, plays a crucial role in increasing bot efficiencies. After bots are implemented and run live, their actions and performance have to be continuously observed to detect exceptions and opportunities for further development or bot redesign. A failing bot can have effects on underlying business processes, e.g. a bot exception can lead to a longer process duration or even an abortion of the whole business process [16]. Therefore, an integrated view of steps performed by bots in the context of existing business processes is needed to analyze these effects and the role of bots in business processes.

On the one hand, leading RPA software can be configured to record logs of the executed steps of bots (bot logs). On the other hand, process mining offers a wide range of tools and techniques to discover process inefficiencies from process logs. Therefore, an integrated analysis using bot and process logs could provide new insights in bot-human interaction, show effects of bots on business processes, show how exceptions of bots are handled and benefit the redesign of bots used in business processes. In this paper, we investigate the following research question: *How can bot logs and process logs be used for process mining to get a better understanding of the behaviors of bots in RPA-enabled business processes?*

To answer this question we first develop an integrated conceptual data model visualized as an ORM diagram describing the relations between bots and business processes. We then specify and instantiate the bot log parser that brings bot logs of the three leading RPA vendors software into the XES format, which is an IEEE standard format for event logs [12]. Moreover, we introduce the log merger that merges bot logs with process logs of the underlying business processes. Next, we propose some process mining measures that help to analyze the merged log. There are many possibilities for new measures, however for this paper we developed two exemplary measures, to illustrate the concept of our approach.

The remainder of this paper is structured as follows: Section 2 introduces a running example. Section 3 summarizes the related work on RPA and process mining. In Section 4 we present the proposed data model, provide a conceptual overview of the approach and develop the bot log parser, the log merger and the exemplary process mining measures. Section 5 concludes the paper and discusses ideas for future work.

2 Running Example

The following exemplary process serves as running example throughout this paper. A visualization of the process can be found in our repository: `https://bit.ly/2Q4CbYr`. Consider an organization with a simplified business process 'Monthly Payroll' which consists of the two activities 'Calculation' and 'Prepare Documents'. Imagine that a bot process 'Auto Calculation', which consists of the three activities 'Open Payroll Spreadsheet', 'Sum up Working Hours' and 'Save and Close Spreadsheet', now automates the so far manual activity 'Calculation'.

Let's assume that process activity 'Calculation' fails in 80% of the cases because bot activity 'Open Payroll Spreadsheet' encounters exceptions when opening the spreadsheet. The failure of 'Calculation' has negative effects on the whole process 'Monthly Payroll'. If we solely analyze the process log, the failures of 'Calculation' and the resulting effects on the rest of the process can be detected. However, the exact reasons of these failures remain unclear. On the other hand, by solely looking at the bot log, the exact reason for the fails, namely the bot activity 'Open Payroll Spreadsheet', can be observed with all relevant variables, however the resulting effects on the business process 'Monthly Payroll' can not. By combining the bot log with the process log, however, the exact causes and the effects of exceptions are observable in an integrated analysis. This exemplary case shows how integrating bot logs in process logs enables new opportunities for process mining and the redesign of bot and business processes.

3 Background

RPA tools mimic human actions on computer systems by interacting with the user interface or by connecting to APIs [3,28]. In the literature, there are various definitions of RPA [14]. RPA is mostly used on rule-based and repetitive tasks and has the potential to replace humans [14,29]. As a result, employees can tackle more complex tasks instead of executing repetitive actions.

Organizations can benefit from using RPA in several ways, e.g. by increasing productivity, by using human resources more effectively as well as a by a more consistent and accurate execution of repetitive tasks by bots [24,26]. Combining RPA with other technologies, such as machine learning, could enable the automation of more complex tasks and provide even more benefits [3,10]. However, there are also challenges when implementing RPA [4,27]. Bots follow the rules written in their code, therefore poorly defined rules can lead to unwanted results [16]. Furthermore, it is important to question the business processes that are automated and to not just blindly automate them, for which an integrated view on bots and the underlying processes is key [16].

UiPath, Automation Anywhere, and Blue Prism are the three leading vendors of RPA solutions [18]. Most RPA tools provide logs that report on bot-executed steps. These bot logs also contain additional payload attributes, such as used input variables and success states. For analytic purposes, some basic bot-related performance measures are provided by RPA tools, e.g. the total number of executions or how many errors occurred [9].

Process mining is based on process logs and allows for discovering as-is models, enhancing existing process models with new insights as well as checking conformance of process enactments [1]. Process logs describe occurrences of historic events and can be extracted from organizations' information systems [5]. By standardizing events along traces, XES is an IEEE format for process logs [12]. The research streams of RPA and process mining already are growing together [20]. Current research predominantly covers the early stages of RPA [15], e.g. by identifying the most suitable processes for RPA [11,22] or desktop activity mining to help constructing bots [21,23]. However, to the best of our knowledge, there is no approach that systematically uses bot logs for applying integrated process mining analysis on automated as well as non-automated process parts.

4 Using Bot and Process Logs for Process Mining

Figure 1 visualizes the conceptual overview of our approach. As a first step we developed a data model that describes the structure and relation of bot processes and business processes along with the required attributes for using bot logs for process mining (Section 4.1). Second, we introduce the bot log parser, bringing bot logs of the three leading RPA vendors into XES format (Section 4.2). Third, we specify the log merger that combines XES-parsed bot logs with process logs to one aggregated merged log (Section 4.3). Fourth, the resulting merged log can be used to gain new insights regarding the role of bots in business processes. For this purpose we suggest exemplary measures as well as develop a concept for visualizing results (Section 4.4). We implemented the bot log parser and the log merger in Java as well as the suggested measures as an extension for the Directly Follows visual Miner (DFvM) [19] in the open-source ProM framework [8]. For more details on this open-source implementation please refer to our repository: `https://svn.win.tue.nl/repos/prom/Packages/BotLogMining/Trunk/src/`

4.1 Conceptual Mapping of Bot Logs and Process Logs

To describe the basic structure of bot logs and develop an approach that enables using them for process mining, we examined the software of the three leading RPA vendors. The tools of all three vendors allow modification of the level of logging, i.e. to which extent bot actions are logged, or to even insert customized logging commands. Therefore, the actual attributes provided in the bot logs can vary from just basic information (e.g., start/end of an action) to detailed payload data (e.g., the accessed URL for bot activity 'Open Browser').

In Figure 2, we describe the structure of bot and business processes along with attributes that are needed to effectively merge bot logs and process logs in Object Role Modeling (ORM) 2 notation [13]. In the following, we describe the key elements of this model:

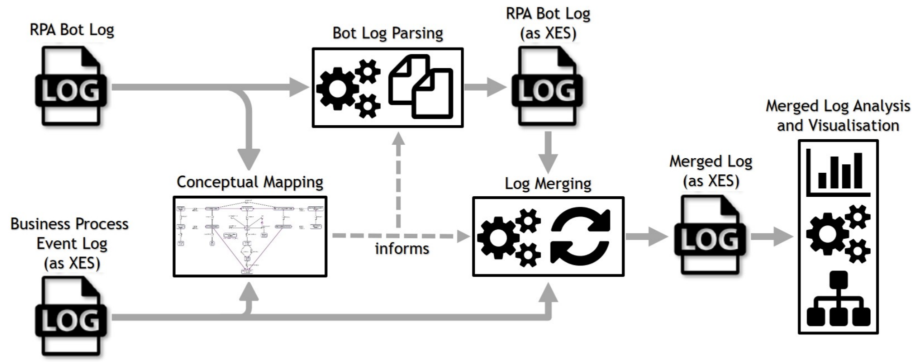

Fig. 1: Conceptual overview of our approach.

- A **bot process** is identified by a name and has a **version**, identified by a number. A bot process (e.g. 'Auto Calculation' in the running example) is an algorithm created with RPA software, including the actions a bot performs.
- A bot process consists of **bot activities**, identified by a name. One bot activity for the bot process 'Auto Calculation' is 'Open Payroll Spreadsheet'.
- An instance of a bot process is a **bot process instance**, executed by a **bot**, with an identifying Id. A bot is a resource and can be allocated to bot processes. A bot process instance is a subtype of a process instance and is identified by a case Id which in turn is executed during a period that consists of a start and an end timestamp.
- A bot process instance consists of **bot activity instances** which are the instances of corresponding bot activities and subtypes of activity instances.
- An activity instance consists of **events** with an identifying Id, recorded at a specific timestamp, going through a lifecycle like 'start' or 'complete', being performed by a bot or not and either failing (because of a reason) or not.

On the right side of Figure 2 a similar structure for business processes as for bot processes is described. In our running example, the business process has the name 'Monthly Payroll' and consists of the business process activities 'Calculation' and 'Prepare Documents'. A bot activity instance is always related to a business process activity instance. In the running example, the instances of the bot activities 'Open Payroll Spreadsheet', 'Sum up Working Hours' and 'Save and Close Spreadsheet' relate to the instances of the business process activity 'Calculation'.

4.2 Bot Log Parsing

To bring bot logs into XES format, relevant attributes have to be extracted from the bot log and standardized according to the specification in Figure 2 to conform to established attribute definitions of the XES standard. Since the bot logs can be customized or the level of logging can be set to different levels,

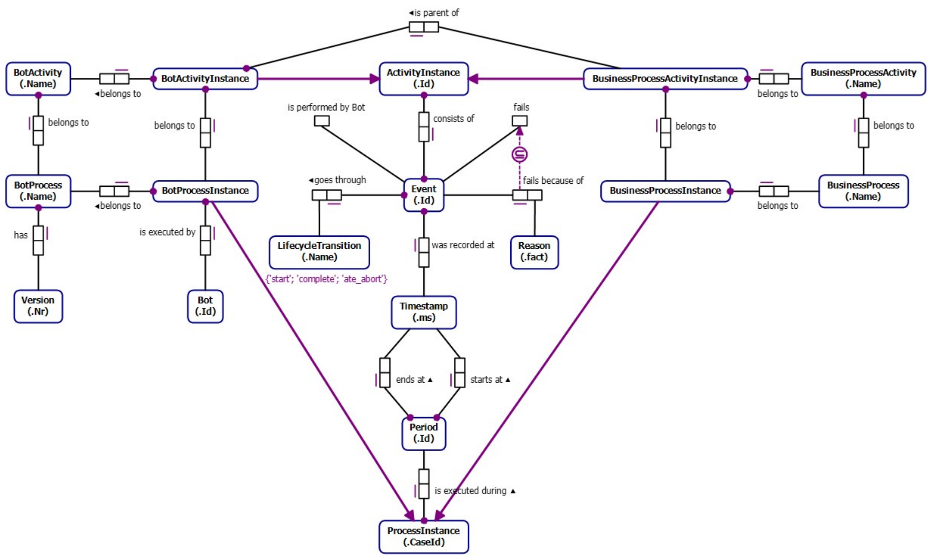

Fig. 2: Structure of bot processes and business processes with relevant attributes.

we need to answer the question, which attributes (as a minimum) should be included in a bot log in order to successfully merge it with process logs.

We define the following standardized attributes for every event in a bot log (i.e. for every action performed by the bot): concept:name, time:timestamp, lifecycle:transition, eventId, caseId, org:resource, botProcessName, botProcessVersionNumber, success and a connecting attribute. Table 1 shows the standardized attributes that can be extracted from various attributes of bot logs, depending on the RPA software used. The notion 'customized' in the table indicates, that the logging can be customized in different ways, depending on the underlying process and therefore the extraction of the standardized attribute can be done in different ways. The connecting attribute will be further explained in Section 4.3.

Table 1: Standardizing bot logs to the XES format.

XES Attribute	UiPath	Blue Prism	Automation Anywhere	Attribute in ORM diagram
concept:name	DisplayName	StageName	*customized*	BotActivityName
time:timestamp	timeStamp	Resource Start+End	first attribute	Timestamp
lifecycle:transition	State	Resource Start+End	*customized*	LifecycleTransition
eventId	fingerprint	StageID	*customized*	EventId
caseId	jobId	*customized*	*customized*	CaseId
org:resource	robotName	*customized*	*customized*	BotId
botProcessName	processName	Process	*customized*	BotProcessName
botProcessVersionNumber	processVersion	*customized*	*customized*	VersionNr
success	State	Result	*customized*	fails
connectingAttribute	*customized*	*customized*	*customized*	*customized*

UiPath provides bot logs in a JSON like format with many different attributes. A figure of the parsing of an exemplary UiPath bot log of the running example can be found in our repository: `https://bit.ly/319wKxM`. The bot log parser extracts the 'Trace Level' logs and uses the provided attributes, according to Table 1, to create a log in XES format. Blue Prism provides logs in a line-by-line format, where several lines can have information about the same action performed by a bot. The bot log parser extracts relevant attributes of an action from different lines and creates a corresponding XES log. In Automation Anywhere the attributes of a bot log can be customized to a high degree. The 'Log-to-File' command can be built into the robotic process algorithm and the exact attributes that are logged for specific bot actions can be defined in this command. According to our conceptual model, we suggest to log at least the attributes provided in Table 1 and we created a corresponding parser for Automation Anywhere bot logs that includes these attributes. However, the code for the parser can easily be adapted to other customized Automation Anywhere bot log structures.

4.3 Log Merging

The goal of the log merger is to create a merged log in XES format by combining a bot log with a process log. As input, the merger takes the process log, the bot log, and the name of the connecting attribute in the process log and in the bot log, respectively. The log merger iterates over the events in the process log and checks the value of the connecting attribute. It then compares this value with the values of the connecting attribute of all events in the bot log. If the values match, the bot event is put at the correct position in the process log, depending on the lifecycle attribute of the process event: If it is in the 'start' lifecycle, the bot event is put after the process event, if it is in the 'complete' lifecycle, the bot event is put before the process event.

In the running example (see Section 2) the instances of the bot activities 'Open Payroll Spreadsheet', 'Sum up Working Hours' and 'Save and Close Spreadsheet' would log a connecting attribute which is also logged by the business process activity instance 'Calculation' (e.g. a common documentId). The 'start' and 'complete' events of the three bot activities would then be placed before the 'complete' event of 'Calculation'. The log merger also puts a new 'bot' attribute to every bot event, which is set to true. This ensures that in the merged log, bot activities can be spotted by looking at this attribute.

The ORM diagram in Figure 2 showed, that every bot activity instance has to belong to a business process activity instance. The connecting attribute includes information about the underlying business process activity the current bot activity belongs to and can vary depending on the use case. It can for example be a document Id, i.e when a bot performs an action the Id of the document the bot is working on is recorded in the bot log. The document Id can then also be found in the process log and thus the connection between bot events and business process events can be observed. If such a connecting attribute is missing, approaches from event correlation could be used to map bot events to corresponding process events (e.g. [25] or [6]).

4.4 Merged Log Analysis and Visualization

A merged log provides opportunities for a more detailed analysis of the underlying processes. The idea is to create new measures and visualizations for process mining that use a merged log as input and provide useful information on the underlying partly automated processes as output. There are many possibilities for new measures, however for this paper we developed two exemplary measures, 'Exception Time Impact' (ETI) and 'Relative Fails' (RF), to illustrate the concept of our approach.

$$ETI_A = \frac{\sum Trace\ rem.Dur.\ (A\,failed)}{\#traces(A\,failed)} - \frac{\sum Trace\ rem.Dur.\ (A\ success)}{\#traces(A\ success)} \tag{1}$$

$$RF_A = \frac{\#events(A\,failed)}{\#events(A)} \tag{2}$$

Equation 1 depicts the ETI measure, that calculates the average impact (in terms of time) which an activity (A) has on the process, if A fails. It compares the average remaining duration of the whole process in cases where A failed to the average remaining duration of the whole process in cases where A did not fail. An interpretation of the measure for the running example could be for example: When the bot activity 'Open Payroll Spreadsheet' fails, the business process 'Monthly Payroll' on average takes 5 hours longer to end. In the visualization of the measure, activities are then colored based on the average time impact in case of failure and based on if they were performed manually or by a bot. This view enables the discovery of critical parts in the process as well as possible effects between bot and human activities. Following the exemplary interpretation of our running example, this could lead to a redesign of the bot process, especially of the bot activity 'Open Payroll Spreadsheet', since this activity on average delays the whole process by 5 hours if it fails.

Equation 2 depicts the RF measure, that calculates the relative exception rate of A by dividing the number of events of A that have the success attribute value 'false' by the total number of occurrences of A in the merged log. In the visualization A is then colored based on the result of that division and based on by whom it was performed. This coloring enables the discovery of often failing activities by bots and humans and possible connections of fail rates of different activities. Further, the fail rates at points of bot-human interaction can be observed and checked for possible patterns.

5 Conclusion and Future Work

In this paper we presented an approach that uses bot logs for process mining, in order to get a better understanding of the behavior of bots in business processes.

We first developed an integrated data model, visualized as an ORM diagram, that describes the structure of bot processes and their relation with business processes. On this basis, we introduced the bot log parser that brings bot logs into the XES format. Furthermore, we introduced the concept of a merged log, and a log merger that combines bot logs with underlying process logs. We then introduced two process mining measures that help analyze merged logs.

We already conducted a first evaluation by parsing a real-life and three artificial bot logs, merging the resulting log with a real-life process log and analyzing the resulting merged log with the two created measures. All datasets and results can be found in our repository. We extend existing knowledge by describing the structure and relation of bot and business processes and by enabling the use of bot logs for process mining.

Our work has some limitations that raise opportunities for future work. First, the basic inputs for the approach are bot logs, which assumes that RPA users have set their logging level accordingly. A more detailed logging can result in more data, which may not always be favored in practice. Second, our approach requires a connecting attribute that allows linking bot actions and business process actions. If there is no such attribute, event correlation techniques have to be applied (e.g. [6] or [25]). Third, more measures are needed to analyze merged logs. We provided two measures to illustrate how new information can be gained from analyzing merged logs. However, there are more opportunities to extract useful information from merged logs, and thus more complex measures are needed.

A possible avenue for future work is to develop an event correlation approach specific for the RPA context. This can help merging bot logs and process logs when connecting attributes are missing, and thus provides further opportunities for generating merged logs. Additional to the two sample measures, more complex measures can be implemented in future work. One idea is to analyze failing bot activities and search for patterns in the corresponding bot log attributes. This measure could be used for mapping exact causes of bot exceptions and exact effects on business processes and thus could benefit bot and process redesign. Moreover, we plan to extend the evaluation with more real-world and artificial data. A further idea is to use more sophisticated techniques like machine learning on merged logs. Bot behavior in new business processes could be predicted and thus possible effects can be derived prior to bot implementation.

References

1. van der Aalst, W.: Process mining: discovery, conformance and enhancement of business processes, vol. 2. Springer (2011)
2. van der Aalst, W., et al.: Process mining manifesto. In: BPM. pp. 169–194 (2011)
3. van der Aalst, W.M., Bichler, M., Heinzl, A.: Robotic process automation (2018)
4. Agostinelli, S., Marrella, A., Mecella, M.: Towards intelligent robotic process automation for bpmers. arXiv preprint arXiv:2001.00804 (2020)
5. Andrews, R., et al.: Quality-informed semi-automated event log generation for process mining. DSS p. 113265 (2020)

6. Bayomie, D., Di Ciccio, C., La Rosa, M., Mendling, J.: A probabilistic approach to event-case correlation for process mining. In: ER. pp. 136–152. Springer (2019)
7. Deloitte: The robots are ready. are you? untapped advantage in your digital workforce (2017), https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf
8. van Dongen, B.F., et al.: The prom framework: A new era in process mining tool support. In: Petri nets. pp. 444–454 (2005)
9. Enríquez, J., et al.: Robotic process automation: A scientific and industrial systematic mapping study. IEEE Access **8**, 39113–39129 (2020)
10. Gao, J., van Zelst, S.J., Lu, X., van der Aalst, W.M.: Automated robotic process automation: A self-learning approach. In: CoopIS. pp. 95–112. Springer (2019)
11. Geyer-Klingeberg, J., Nakladal, J., Baldauf, F., Veit, F.: Process mining and robotic process automation: A perfect match. In: BPM (Dissertation/Demos/Industry). pp. 124–131 (2018)
12. Günther, C., Verbeek, H.: XES v2.0 (2014), `http://www.xes-standard.org/`
13. Halpin, T.: Orm 2 graphical notation. Technical Report ORM2–02 (2005)
14. Ivančić, L., Vugec, D.S., Vukšić, V.B.: Robotic process automation: Systematic literature review. In: BPM. pp. 280–295 (2019)
15. Jimenez-Ramirez, A., et al.: A method to improve the early stages of the robotic process automation lifecycle. In: Caise. pp. 446–461 (2019)
16. Kirchmer, M., Franz, P.: Value-driven robotic process automation (rpa). In: BMSD. pp. 31–46. Springer (2019)
17. Lacity, M., Willcocks, L.P., Craig, A.: Robotic process automation at telefonica o2 (2015)
18. Le Clair, C., UiPath, A.A., Prism, B.: The forrester wave^{TM}: Robotic process automation, q2 2018. Forrester Research (2018)
19. Leemans, S.J., Poppe, E., Wynn, M.T.: Directly follows-based process mining: Exploration & a case study. In: ICPM. pp. 25–32 (2019)
20. Leno, V., Polyvyanyy, A., Dumas, M., La Rosa, M., Maggi, F.M.: Robotic process mining: Vision and challenges. BISE pp. 1–14 (2020)
21. Leno, V., et al.: Action logger: Enabling process mining for robotic process automation. In: BPM demos. pp. 124–128 (2019)
22. Leopold, H., van der Aa, H., Reijers, H.A.: Identifying candidate tasks for robotic process automation in textual process descriptions. In: EBPISM, pp. 67–81 (2018)
23. Linn, C., Zimmermann, P., Werth, D.: Desktop activity mining-a new level of detail in mining business processes. In: APSN workshops (2018)
24. Madakam, S., Holmukhe, R.M., Jaiswal, D.K.: The future digital work force: robotic process automation (rpa). JISTEM **16** (2019)
25. Pourmirza, S., et al.: Correlation miner: mining business process models and event correlations without case identifiers. IJCIS **26**(02), 1742002 (2017)
26. Slaby, J.R.: Robotic automation emerges as a threat to traditional low-cost outsourcing. HfS Research Ltd **1**(1), 3–3 (2012)
27. Syed, R., et al.: Robotic process automation: Contemporary themes and challenges. CI **115**, 103162 (2020)
28. Tornbohm, C., Dunie, R.: Gartner market guide for robotic process automation software. Report G00319864. Gartner (2017)
29. Willcocks, L.P., Lacity, M., Craig, A.: Robotic process automation at xchanging (2015)