A method for objective performance benchmarking of teams with process mining and DEA

A METHOD FOR OBJECTIVE PERFORMANCE BENCHMARKING OF TEAMS WITH PROCESS MINING AND DEA

Complete Research

Aysolmaz, Banu, Eindhoven University of Technology, Netherlands,
b.e.aysolmaz@tue.nl
Nemeth, Marcel, Maastricht University, Netherlands,
m.nemeth@alumni.maastrichtuniversity.nl
Iren, Deniz, Center for Actionable Research, Open Universiteit, Heerlen, Netherlands,
deniz.iren@ou.nl

Abstract

Process performance is influenced by the teams involved in those processes. However, objectively evaluating team performance is challenging. Process mining literature has mostly focused on the characteristics of social networks rather than their performance. Following design science research principles, we designed a method for objectively benchmarking team performance using process mining, data envelopment analysis, and social network analysis. Our stepwise method provides guidelines on choosing appropriate data, extracting teams, identifying and calculating relevant metrics, benchmarking team performance, and determining improvement areas by considering the process context. We demonstrate the application on a real-life loan management process. We conducted interviews with business professionals to validate the usefulness and applicability. The results show that the method can provide an objective and transparent mechanism for comparing the performance of teams. The method can be used by organizations to form efficient teams and facilitate an increase in productivity.

Keywords: team performance, performance benchmarking, process mining, data envelopment analysis, organizational social networks.

1 Introduction

Business processes, ranging from manufacturing to supply-chain management and marketing, constitute the heart of business operations (Provost and Fawcett, 2013). Processes, people, and technologies are building blocks of an organization (Dumas et al., 2018). Among these, the human is seen as the most crucial element for successful operation of a business (Thevendran and Mawdesley, 2004). However, it is not only the knowledge and skills of individual employees that affect organizational performance, but also multiple employees working together for a collective purpose (Newell, Tansley, and J. Huang, 2004).

Employees working on a certain business process together establish a social network (Song and Aalst, 2008). Social networks, or teams, play a crucial role in the way organizations are run and problems are solved (Yang and Tang, 2004). Organizational performance is closely related to how teams are composed (Sparrowe et al., 2001). More specifically, process performance is found to be significantly influenced by teams and their performances (Furht, 2010; Kratzer, Leenders, and Van Engelen, 2010). Since processes are usually performed collaboratively, process performance may change based on the team structure formed to execute those processes (Kumar, Dijkman, and Song, 2013b). For example, in one process instance, employee A may hand over a document to B, who hands it over to D. In a different case, the team composition for performing the process instance may be as [A-D-E]. The team [A-B-D] might perform better than the team [A-D-E], although both teams execute the same activities and share two identical members. The variation in performance of an individual may be caused by their affiliation with one team rather than another (Manski, 2000). Thus, observing and evaluating individuals in an isolated manner might not suffice to evaluate and improve process performance. Focusing on the performance based on the composition of teams may present an opportunity to improve process performance. However, it is a challenge to objectively compare how different teams perform so that they can be set up effectively (Levy and Williams, 2004). Previous studies have mostly used perceptual or indirect measurement methods to evaluate and compare team performance (Franco-Santos, Lucianetti, and Bourne, 2012).

Process mining is a data science technique that can provide an objective and accurate view of business processes based on increasingly-available event logs from process-aware information systems (Aalst, 2016). Process mining has been used to provide various insights on organizational aspects of processes (W. Zhao and X. Zhao, 2014) such as revealing the organizational structure (Aalst and Song, 2004; Song and Aalst, 2008), collaborations (Schonig et al., 2018) and behavioral patterns (Delcoucq et al., 2020). Team dimension has been considered for work assignment and providing descriptive information (Schonig et al., 2018). In summary, current process mining studies on organizational aspects of processes have mostly focused on the characteristics of the social networks and studied basic process performance measures as descriptives, although it is suggested that team analysis on processes should be extended with performance data (Song and Aalst, 2008). Resource performance from event logs has been studied in an individual level by Pika et al. (2017), who benchmarked the performance of human resources by the tasks they performed. Their research set an example for benchmarking resource performance as a way of process performance management. Thus, there seems to be a potential to use event logs and process mining to objectively assess performances of teams working on processes and improve process performance.

In this paper, we propose a method to objectively benchmark the performance of teams that execute processes, identify improvement areas, and analyze team compositions to deduce insights on performance. The method implements process mining to automatically extract teams and calculate related metrics, and Data Envelopment Analysis (DEA), a method for benchmarking the efficiency of units based on a non-parametric function, to compare team performances using a set of metrics calculated through event logs. Even when objective data such as event logs are used, it is still challenging to choose appropriate data and identify proper analysis techniques for sound and reliable performance evaluation (Behn, 2003). These decisions would have a high impact on the measured performance. Our method includes tasks and guidelines so that an organization can consider different aspects of their business context and business process to select relevant data from the event log, metrics, and performance analysis techniques. We have followed a design science research methodology (Gregor and Hevner, 2013) to guide our research process to develop the method. We used CRISP-DM, a widely-accepted data mining process methodology, to structure the steps of the method (Wirth and Hipp, 2000). We demonstrate the application of the method on a real-life event log available as a research data set and evaluate it via interviews with professionals to validate the usefulness and applicability. Our study extends research on process mining in combination with performance measurement and Social Network Analysis (SNA). In this way, the method may provide a more accurate and objective performance comparison of team of employees working together to execute processes and help to identify improvement areas. It may open up new opportunities for managers to identify teams and assign individuals for process execution.

The remainder of this paper is structured as follows. Section 2 presents a brief overview of the state of the art on process mining related to organizational aspects, performance management, and performance benchmarking techniques. Section 3 explains our design science research methodology. Section 4 describes the proposed method in a stepwise manner. In Section 5, we present the demonstration and evaluation of our method and results. In Section 6, we highlight the implications, conclude the study, and provide directions for future work.

2 Background

In this section, we first discuss the related work on the uses of process mining on the organizational aspects of processes (Section 2.1) and relevant studies on team performance management (Section 2.2). We then introduce the background on the use of benchmarking for team performance and DEA as a performance benchmarking method (Section 2.3).

2.1 Process mining for organizational aspects

Social networks are critical for the success of an organization, since they identify how individuals work together to unite different levels of experience and skills for a common goal (Lin, 2001; Molm, 1994). Process mining can be used to extract information related to the organizational structure and social networks, in other words teams, related to processes based on objective evidence found in event logs (Aalst, 2016; W. Zhao and X. Zhao, 2014). A social network consists of individuals and relations among them, which is typically visualized as a graph with nodes being the individuals and connections between the nodes representing the relation between the individuals. Social networks related to processes have been identified based on various relations such as joint process cases, joint activities, and handover of works (Aalst, Reijers, and Song, 2005; Aalst and Song, 2004). The structure of social networks has been analyzed to understand the characteristics of social networks, such as the density of the network and betweenness, closeness, and power of individuals in the network (Aalst, Reijers, and Song, 2005) and discover organizational models (Hanachi, Gaaloul, and Mondi, 2012; Song and Aalst, 2008). Resource-related information in the event logs has further been used to discover resource behaviors such as competence and constraints on how to assign resources to activities (Z. Huang, Lu, and Duan, 2012; Schönig et al., 2015), mine how multiple resources collaborate on activities (Kumar, Dijkman, and Song, 2013b; Schonig et al., 2018), and find out individuals with similar behavioral patterns (Delcoucq et al., 2020). Overall, process mining literature establishes event logs as a relevant and objective source to reveal teams that execute processes and provide necessary tools to analyze the team compositions.

2.2 Team performance management

Performance management enables organizations to identify malfunctioning points and improvement opportunities in processes (Behn, 2003). The studies in the literature has mostly offered performance measurement approaches that rely on managers' personal observations and subjective judgments on employees and processes (Franco-Santos, Lucianetti, and Bourne, 2012). They look at the behavior from outside and do not grasp all relevant aspects of performance (Neely, Gregory, and Platts, 2005). Process mining techniques, on the contrary, view processes from an internal perspective and provide objective evidence (Z. Huang, Lu, and Duan, 2012). Moreover, relevant data can be gathered easily through event logs generated automatically by information systems (Z. Huang, Lu, and Duan, 2012), unlike subjective approaches that involve manual data collection (Carrington, Scott, and Wasserman, 2005). Process mining has been used to evaluate performance of individuals in organizations. Pika et al. (2017) suggested an approach to measure the productivity of individuals and compare it with others. Swennen et al. (2015) suggested a number of process metrics such as resource productivity.

Managing team performance has been found as important as individual performance (Sparrowe et al., 2001). Studies have investigated if certain characteristics of social networks can be associated with performance. For example, the degree of connectedness, i.e., the number of ties among employees, was found to be indirectly related to their satisfaction and performance (Brass, 1981). Among the studies on using process mining to extract organizational models and social networks, team performance is tackled only through the identification of bottlenecks between handovers (Kumar, Dijkman, and Song, 2013b; Song and Aalst, 2008), analysis of workload differences for individuals in a team (Z. Huang, Lu, and Duan, 2012; Nakatumba and Aalst, 2010), and reporting of average process durations for teams (Schonig et al., 2018). Therefore, although performance is seen as an important aspect of teams in a process context (Song and Aalst, 2008), it appears that the performance of individuals that work together as a team to execute processes has not been closely examined.

2.3 Performance benchmarking and DEA

Benchmarking is defined as the systematic comparison of the performance of different Decision Making Units (DMU), such as firms, departments, teams, or individuals (Bogetoft and Otto, 2011). Frontier analysis is a preferred benchmarking approach, which identifies the highest-performing (i.e., frontier) DMU that produces the maximum output from a given input (Bogetoft and Otto, 2011). Among other frontier analysis methods, DEA has been found to be flexible regarding the underlying data and assumptions (Cooper, Seiford, and Zhu, 2011). DEA does not require information about how the production process operates to transform inputs to outputs. The inputs and outputs are usually work, money, raw materials, or production units, but can also be human resources and productivity (Bogetoft and Otto, 2011). Linear programming is used to estimate the optimal combination of weighted inputs and outputs a DMU can produce (Cooper, Seiford, and Zhu, 2011). The DMU, or a set of DMUs, that fulfill this assumption is called the best practice or efficient frontier. All remaining DMUs are considered suboptimal and compared to the frontier by calculating a relative efficiency score, which is between 0 and 1 (Pika et al., 2017).

DEA has previously been used in business process performance evaluation research. Burger and Moormann (2008) demonstrated the application of DEA on a process level. Koch-Rogge et al. (2014) and Dugelova and Strenitzerova (2015) showed that DEA could be used to evaluate resource behavior in general. Pika et al. (2017) combined both approaches to evaluate the performance of individuals based on event log data from a process. Building on these studies, we use DEA for the performance evaluation of teams in organizations formed by executed business processes.

3 Research design

To develop the proposed method, we followed the design science research paradigm (Gregor and Hevner, 2013) and implemented the four research steps proposed by Peffers et al. (2007). We performed our research in four steps; problem identification and motivation, definition of the objective of the solution, design and development, and demonstration and evaluation.

In the first step, we *identified the problem* as the objective evaluation of the performance of the teams that work on certain processes together in organizations. Although performance management of teams has been shown important to reveal potential improvements and impact overall organizational performance, in the current literature, objective performance comparison of process-related teams through process mining has not been particularly studied. Process mining literature, on the other hand, has established the importance of analysing process-related teams and conducted research on objective performance evaluation, albeit on individual level. Our *motivation* is to fill this gap by developing a method to objectively benchmark teams for their performances using event logs. Next, to be able to use process mining and other relevant data analysis methods on event logs (such as DEA), we defined the *objective of our method* as to specify the tasks, inputs, and outputs and provide guidelines on how the tasks can be executed as a data mining process. To achieve this objective, we have designed the method as a data mining process, following the common steps of suggested data mining process methodologies (Plotnikova, Dumas, and Milani, 2020; Saltz et al., 2018). We have *designed the method* in a stepwise manner and *developed* software codes to implement relevant parts, as we introduce in the next section. As the last step, we *demonstrated* the application of the method on an example real-life loan management process. We *evaluated* the usefulness and limitations of the method through interviews.

4 Method for team performance benchmarking

In this section, we describe the steps of our method to benchmark the performance of teams executing business processes, using event logs. We use CRISP-DM (Wirth and Hipp, 2000) as the guiding framework to design our method, a widely accepted methodology as a standard process model for data mining (Aysolmaz, Dau, and Iren, 2020; Plotnikova, Dumas, and Milani, 2020), since our suggested solution involves process mining and other data analysis methods. Our method is composed of five steps and takes as input the event log file and knowledge of the investigated process, as depicted in Figure 1. The first step is (1) *Data selection* to choose event logs of relevant cases and get a list of teams. This corresponds to data understanding and data preparation phases of CRISP-DM. Next steps are (2) *Identification of metrics* that are relevant to the business process context and calculation of metric values for teams and (3) *DEA model implementation* to obtain efficiency scores of teams and identify improvement areas for sub-optimal teams. These align with the phases of modeling and evaluation of alternative solutions of CRISP-DM, since these steps require the selection and evaluation of metrics and DEA settings. Next, two steps are performed in conformance with the deployment phase of CRISP-DM: (4) *Evaluation of individual effect on team performance* to reveal how individuals may affect team performance, and (5) *Evaluation of the effect of team compositions on team performances* by investigating the similarity among teams and analyzing the structure of social networks. Each step is explained in the sections below. We do not provide formal definitions of the concepts in this study, such as the derivation of case-based teams from an event log, some of which can be found in other studies (Song and Aalst, 2008). We rather focus on providing explanations and guidance to consider the business context. We developed Python and R scripts to execute the method from step 1 (partially) to step 5[1].

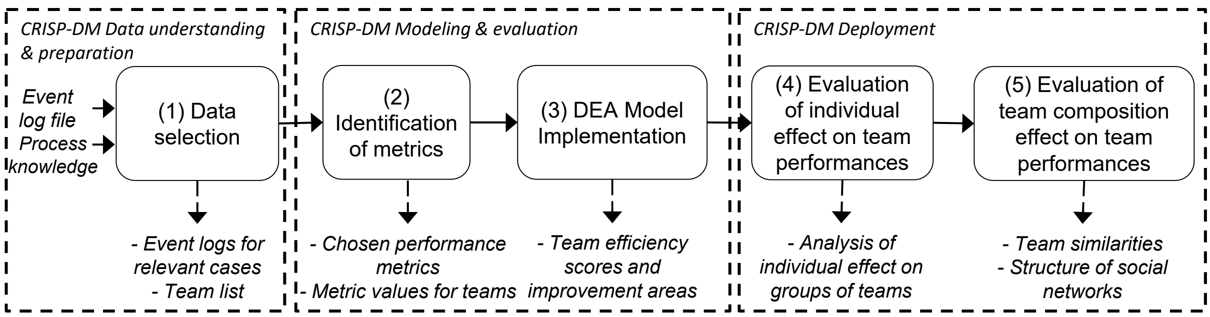

Figure 1: Overview of the method

4.1 Step 1- Data selection

The first step includes activities to identify data quality issues, clean data, and select data relevant for team performance evaluation, which are typical activities in data and process mining tasks (Aalst, 2016; Wirth and Hipp, 2000). The goal is to create event logs for relevant cases and automatically extract the list of

[1] The codes together with their example application for each step as described in Section 5 can be found at: https://doi.org/10.5281/zenodo.4673072

teams. The requirement for this step is making the decisions regarding incomplete and irrelevant cases based on process knowledge. Irrelevant cases in the data would lead to erronous identification of teams and calculation of metrics. The following list of guidelines should be considered to select relevant cases.
- Remove cases that have not been properly completed (Aalst, 2016).
- Remove invalid cases with missing data (Wirth and Hipp, 2000), such as cases including activities without any human resource (henceforth referred to as resource) assigned and with zero processing time.
- Consider removing process variants that may not be comparable in performance (Aysolmaz et al., 2019), such as different cars in the same production process.
- Consider removing cases that complete in unusual ways (Aalst, 2016) due to external factors (e.g., customer not paying for an accepted product), since they may prevent a fair team comparison.
- Consider dropping cases that include infrequent activities and individuals. Such cases may show rare diversions from the process flow handled by specialized resources (Rodrigues et al., 2017). Thus, they may not provide a good basis for performance comparison.
- Consider dropping outlier cases in terms of duration (too fast or too slow) or number of activities, which can interfere with evaluation.
- Remove cases that have only one resource, since those cases do not contain any team.

After cleaning and filtering of the event log based on the decisions, this step includes automated extraction of *team list* from the event logs of relevant cases. For each case, the set of resources that work together on that case is extracted. In this way, the list of teams for all cases in the log are built. It can be considered to remove the cases related to the teams that performed too few cases together (Schonig et al., 2018), to establish reliable data for team performance.

4.2 Step 2- Identification of metrics

The goal of this step is the selection of the metrics relevant for performance evaluation and automated calculation of the metric values for teams. The requirement for this step is making the decisions on the relevant metrics based on process knowledge. If metrics are not chosen appropriately, the calculated metric values and team performances would not align with the business value created in the process. The following guidelines should be considered to make the decisions and execute this step.
- *Choose relevant metrics* based on the process knowledge (Neely, Gregory, and Platts, 2005). Through the use of DEA, our method allows the flexibility to incorporate any metric set that is relevant to the context. Although it is advised not to define too many metrics to avoid unnecessary complexity (Franceschini, Galetto, and Maisano, 2007), there is also no limitation on the number of metrics for the method. The example metrics and considerations below can be used as a guideline for this decision. In Table 1, we present several *performance-related metrics* from literature. These metrics are adapted to a case level to provide a benchmark for the teams performing cases to produce a specific output together. The suggested metrics measure the performance in different dimensions, namely, productivity, skills, and utilization, to allow for general inferences about process performance (Franceschini, Galetto, and Maisano, 2007). To calculate the chosen metrics, it may be necessary to make assumptions based on the available data and process knowledge. For example, if *processing time* is chosen, the event log should include start and end times for activities. If there are activities that require input from an external party, it may be considered to remove the duration of those activities from the time metrics so that it does not obscure the performance of the teams. Business value can vary widely and a higher monetary amount can indicate higher value in one process, e.g., orders, and vice versa in another, e.g., purchases. Activity completion can be seen as an indication of high performance, since more activities are executed in the case as needed. Conversely, a lower number of activities can be seen as an indication of higher efficiency and better performance indicating that the team can achieve results with less activity. Such contextual information would identify the setting of input and output metrics of DEA, as described in the next section. If activities require different levels of effort, a weight can be assigned to activities, which can then be used in metric calculations (Pika et al., 2017).

Dimension	Metric	Description	Unit	Source
Productivity	Processing time	Total execution time of all activities a case, excluding idle time	Hours	(Nakatumba and Aalst, 2010) (Pika et al., 2017) (Swennen et al., 2015)
	Throughput time	Execution time of a case including idle times between activities	Hours	
	Business value	Value of the output created in the case, e.g., purchase amount	Euros	
	Activity duration	Average duration of all activities completed per resource in the case normalized for all resources	Hours	
Utilization	Availability	Total availability time of all resources during the execution of the case	Hours	(Pika et al., 2017)
	Activity completion	Total number of activities completed by all resources in the case	Number	
Skill	Distinct activities	Number of distinct activities performed by a single resource per case summed for all resources in the case	Number	

Table 1: Performance metrics from literature for team performance measurement

- *Calculate metric values for teams* automatically for the teams in the team list. When a team performs more than one process case together, the calculated metrics are normalized to team level by taking an average of the values over cases.

4.3 Step 3- DEA model implementation

The goal of this step is the implementation of DEA to calculate team efficiency scores and improvement areas for the sub-optimal teams. The requirement is the decisions for the DEA model configuration. If the model is not configured properly, the team efficiency scores would not be adjusted based on the business context and priorities. The following guidelines should be considered to implement the DEA model.

- Specify the previously chosen metrics as *inputs or outputs for the DEA model*. The metrics that are aimed to be low in the process are set as input parameters, such as processing and throughput time. Output metrics should be chosen as those that need to be maximized for higher performance.
- Determine the objective function of DEA, i.e., *minimize inputs (input-oriented) or maximize outputs (output-oriented)*, based on the priorities of the organization (Islam and Manzoni, 2009).
- Identify the *return-to-scale (RTS) parameter*, i.e., the setting of how inputs or outputs behave when one or the other is changed. Constant RTS (CRTS) is used when, for example, doubling the input results double the output, while increasing RTS (IRTS) is for cases that the output increases in greater proportion than the input. Decreasing RTS (DRTS) is used when the output increases less than the increase in input. RTS should be set carefully to make sure that it reflects the behavior of the input and output variables in the business context (Pika et al., 2017).
- *Run the DEA algorithm* to automatically calculate the team efficiency scores and improvement areas. The outcome of DEA is a benchmark table where each row shows an individual team, or DMU in DEA terminology, and a team *efficiency score*, where the most efficient teams are assigned a score of 100%. The remaining teams are assigned lower efficiency scores with respect to the most efficient ones. These indicate unexploited improvement areas. Such DMUs have one or more benchmark DMUs that can be used to identify improvement areas. Generally, a benchmark DMU achieves a higher efficiency with the same optimal weights as the DMU that is being examined. For such DMUs, the model gives a weight score, *lambda*. By multiplying lambda scores with the metrics, the improvement areas that can be achieved by a sub-optimal team are identified. The application of these calculations are exemplified in the evaluation (specifically Section 5.1.3). The organization can decide which of the improvement areas they can implement based on their business goals and constraints.

Upon revealing the differences in team performances, the organization can perform further analysis to understand what causes the teams to deviate in performance (Furht, 2010). One possible cause can be the effect of individuals and another can be the composition of teams. Our method provides guidance on further analyses about these causes in the following steps.

4.4 Step 4- Evaluation of individual effect on team performances

The goal of this step is to investigate if there are individuals that affect team performances in all teams or a group of teams with a certain performance profile (e.g., top-performing or low-performing teams based

on team efficiency scores). The requirement for this step is that resources should be working in multiple teams. The step can be executed automatically.
- *Investigate the number of teams that the resources worked in:* This gives an indication of if individuals engage in work with different teams or participate in cases mostly with similar teams.
- *Analyze per resource the mean efficiency of teams the resource worked in:* This analysis may give a picture of resources that work more in high (or low) performing teams. We name this as *mean team performance of resources*. This is an indication of not individual capabilities but the performance of teams that the individual takes part in.
- *Investigate the teams of resources with different team performance profiles:* High-ranked individuals are examined for team performance and checked if they take part mostly in top-performing teams. This may give an indication if certain individuals mostly cause higher team performance or if this is not a related factor. The analysis is also performed to investigate the resources engaged in low-performing teams.

4.5 Step 5- Evaluation of team composition effect on team performances

The goal of this step is to investigate the differences in team compositions and reveal possible reasons behind it. There is no requirement for this step, it can be executed automatically. The outputs are used to perform the analysis as guided in the following steps.
- *Examine the similarity of teams with a certain performance profile:* The similarity of team compositions are calculated for every pair of teams that fit a certain performance profile, e.g., frontier teams. A higher similarity among the teams in a performance profile group may indicate a stronger effect of team composition on team performance. We use cosine similarity, a similarity metric widely-used in various domains (Bayardo, Ma, and Srikant, 2007) to calcuate similarity of team compositions. To calculate the cosine similarity, each team is represented as an n-dimensional vector, where n is the number of all individual resources that takes part in one or more teams. The values in team vectors are binary; 1 indicates that the resource is in the team, and 0 means it is not.
- *Investigate ties among resources in teams with a certain performance profile:* Social network graphs are developed and social network metrics are calculated automatically for this analysis. A node in the graph represents a resource and a connection among the two resources is established by “participating in the same team". Density metric is used to compare the structure of a social network, which measures the number of social ties in the network per all possible ties (Furht, 2010). Comparing the structure of the social networks of group of teams may give hints about the underlying reasons of their performance difference (Furht, 2010). For example, if top-performing teams are found to have similar team compositions, it can be considered that the resources of those teams work together more often, thus, develop an affinity towards each other that positively affects team performance.
- *Compare the density of different groups of teams:* Differences in the social network characteristics of group of teams may point out to differences in the structure of the social networks, which helps to reveal the underlying reasons in team performance differences.

5 Evaluation and results

In this section, in conformance with the last step of the design science research (Peffers et al., 2007), we first demonstrate the application of our method on a real-life loan management process available as a research data set, then, we present the results of the interviews to evaluate the usefulness and limitations of the method in organizational settings.

5.1 Application on Real-Life Process Data

We applied our method to measure the performance of teams that perform a business process in an organization following the steps described in Section 4[2]. We used an anonymized real-life event log

[2] The execution of the method on the event log and relevant output files can be found in the code package.

Case ID	Activity	Resource	Start Time	Complete Time	Loan Goal	Offr Amnt
1226577675	A_Create Application	User_1	1/5/16 8:27:37PM	1/5/16 8:27:37PM	Car	
1226577675	W_Complete application	User_108	1/6/16 7:32:30PM	1/6/16 7:43:15PM	Car	
1226577675	O_Create Offer	User_108	1/6/16 7:42:51PM	1/6/16 7:42:51PM	Car	10000
1226577675	O_Sent (mail&online)	User_108	1/6/16 7:43:15PM	1/6/16 7:43:15PM	Car	
1226577675	W_Call after offers	User_108	1/6/16 7:43:15PM	1/6/16 7:43:15PM	Car	
1226577675	W_Validate application	User_30	1/13/16 3:19:13PM	1/14/16 11:18:09AM	Car	
1226577675	O_Accepted	User_30	1/14/16 11:18:09AM	1/14/16 11:18:09AM	Car	

Table 2: An excerpt from the event log used in the application of the method

obtained from 4TU.Centre of Research Data (Van Dongen, B.F., 2017). This log represents a loan management process from a Dutch financial institution containing events from 1 January 2016 until 1 February 2017. We chose this event log for a number of reasons. First, it represents a real-life process, which is valuable to show the applicability, benefits, and limitations of the method. Second, our evaluation is reproducible since the data is publicly available. Third, the log contains rich information important for evaluating the performance, which may not always exist in all event logs, such as start and end timestamps, resources, and value of loan offers. Table 2 shows an excerpt from the log file for an exemplary simplified process case. *Case ID* is the unique identifier assigned to each process case, i.e., a loan application. Activities that start with *A* and *O* indicate the status changes of a loan application and offer respectively, thus, they are instant. *Resource* indicates the user that performed the activity. *Loan goal* indicates the type of the loan such as car, mortgage, or home renovation. *Offr Amnt* is the value of the loan offer.

5.1.1 Step 1- Data selection

The event log contained 31.509 cases in total. We filtered the data to select relevant cases from the event log based on an understanding of the process knowledge in previous studies (Rodrigues et al., 2017). We first removed cases that are unfinished (cases that did not reach a final decision activity) and invalid (cases having zero processing time in total). This resulted in the removal of 6688 cases from the log. Then, we found cases with infrequent activities ($<$1% frequency) and with roles attending few cases ($<$1% cases). We assumed that process variants for different loan goals (i.e., loan for home or car) are comparable to each other for performance. Next, we kept cases that are only finalized with a successful loan and removed the rest. We removed outlier cases that have exceptionally few or many number of activities (5%). Our event log for relevant cases eventually contained 7238 cases and 93 unique resources[3]. By running the script, we automatically extracted the team list. We removed the resource User_1 from the teams, since it is assumed to be an automation system rather than a human resource. We confirmed that we have multiple process cases per team. This resulted in 237 teams and 81 resources having performed more than one case.

5.1.2 Step 2- Identification of metrics

Based on our knowledge of the loan management process, we chose four metrics among the ones in Table 1 to evaluate the performance of the teams working in the process: throughput time and processing time, and loan amount (business value) as productivity metrics and activity completion as a utilization metric. The selection of metrics was done based on the authors' own judgment of the process. Both the set and number of metrics could be different, which may change the results but not the application of the method. We ran our script to automatically get metric values for the teams in the team list. All metrics were normalized over the cases that the team worked on. We assumed that all activities require equal effort, so we did not assign any weight to them. This decision affects the metric values and eventually the comparison of the team performances.

5.1.3 Step 3- DEA model implementation

We configured our DEA model in conformance with our metric choices. Processing and throughput times are assigned as input parameters, as they would be preferred to be low. Activity completion and loan

[3] We used Disco tool by Fluxicon to filter and export the data to csv format up to this step.

			Weights (lambda) for benchmark team			
Team ID	**Resources**	**Efficiency score**	**4132**	**4519**	**5941**	**6386**
5941	14, 75, 119, 125, 133	100.0%	0	0	0	0
4985	27, 61, 90, 121	94%	0	0	0.71	0.29
3536	5, 109, 121	71%	0	0.26	0.37	0.15
784	25, 115, 117	48%	0.31	0.02	0.40	0

Table 3: DEA benchmarking results for four teams and their weights

Team	**Team 5941** (lambda: 0.71)		**Team 6386** (lambda: 0.29)			**Team 4985** (benchmarked team)	
Metric	Value	Value*Lambda	Value	Value*Lambda	Totals	Value	Excess Use
Input metrics							
Processing time (mins)	9.7	6.89	3.5	1.02	7.91	8.5	0.6
Throughput time (hrs)	213	151	194	56	207	678	471
Output metrics							
Loan amount	27500	19525	12500	3625	23150	15250	-7900
Activity completion	22	15,6	15	4,4	20	20	0

Table 4: Example improvement area calculation for Team 4985

amount are set as output parameters, since we assume that the higher values for them indicate higher performance. We set our model as input-oriented, indicating that we want to minimize the inputs, together with the RTS parameter set as decreasing. The resulting DEA benchmark table calculated automatically contains the efficiency scores of 237 teams. 10 teams at the top of the list received a 100% efficiency score, meaning that they performed the highest and equally well among all teams (which does not necessarily indicate equal values of metrics). Table 3 shows an excerpt from the benchmark table generated by the DEA model. The first team, 5941, achieved an efficiency score of 100%, i.e., they used their inputs in the most efficient way to produce the respective outputs, together with 10 other teams. The rest of the teams have lower efficiency scores down to 22% in the dataset. For example, team 4985 has an efficiency of 94%, which can be benchmarked with top-performing teams of 5941 and 6386 with lambda values of 0.71 and 0.29 respectively as provided by DEA. Lambda values are weight scores identified for each sub-optimal team with respect to one or more reference frontier team. They are used to calculate how the performance of the sub-optimal team can be improved with respect to the reference frontier teams. An example calculation is depicted for team 4985 in Table 4. Team 4985 has 471 hours of more throughput time and 0.6 minutes of more processing time. This results in 7900€ of less value achieved with the same activity completions. These gaps would have to be closed to improve the efficiency of team 4985 to 100%.

5.1.4 Step 4- Evaluation of individual effect on team performance

We first investigate the number of teams the resources worked in, which changes from 1 to 41 with a mean of 9.2 and standard deviation of 9.6. This shows a lot of diversity in how resources participate in team formation[4]. Next, *mean team performance of resources* is calculated by averaging, for each resource, the efficiency scores of the teams that the resource worked in. This value deviates from 0.40 to 1.0 with a mean of 0.6 and standard deviation of 0.12. High-ranked individuals are identified as the top 10% of all resources, which amounts to 8 resources, based on the ranking for the mean team performance of resources. We examined how many of these 8 resources work in the 10 top-performing teams. Surprisingly, only 3 individuals out of 8 actually worked in these teams. This point may indicate for this process that the team performance is more dependent on the synergy of the team than the performance of its members.

5.1.5 Step 5- Evaluation of team composition effect on team performance

To compare the similarity of teams with a certain performance profile, we calculated the cosine similarity within five different groups of teams; top-10 (100% efficiency), bottom-10, top-20 (>90% efficiency), and bottom-20 teams in performance, and all 237 teams, as shown in Table 5. The top-10 teams have a lower similarity score (0.067) than bottom-10 teams (0.112) and all teams (0.077). Thus, bottom-10 teams are more similar to each other than top-10 teams in terms of their team composition. The analysis of social

[4] A dot plot of individuals working in teams can be found in the code package.

	#Pairwise comparisons	Cosine similarity	Density
Top-10	45	0.067	0.158
Top-20	190	0.068	0.110
Bottom-10	45	0.112	0.147
Bottom-20	190	0.086	0.116
All 237	28680	0.077	0.164

Table 5: Cosine similarity and density values among five groups of teams

networks, which are formed based on the resources working in the same team, provides more information on the structure of these groups. Table 5 reports the density of different groups and Figure 2 shows social network graphs of top and bottom-10 teams. Top and bottom-10 teams are denser than top and bottom-20, indicating that these teams have stronger social ties within. The investigation of the graphs reveals that both groups have a distinct set of members that do not take place in other teams. These results need to be investigated in the business context to be able to make decisions about future team compositions.

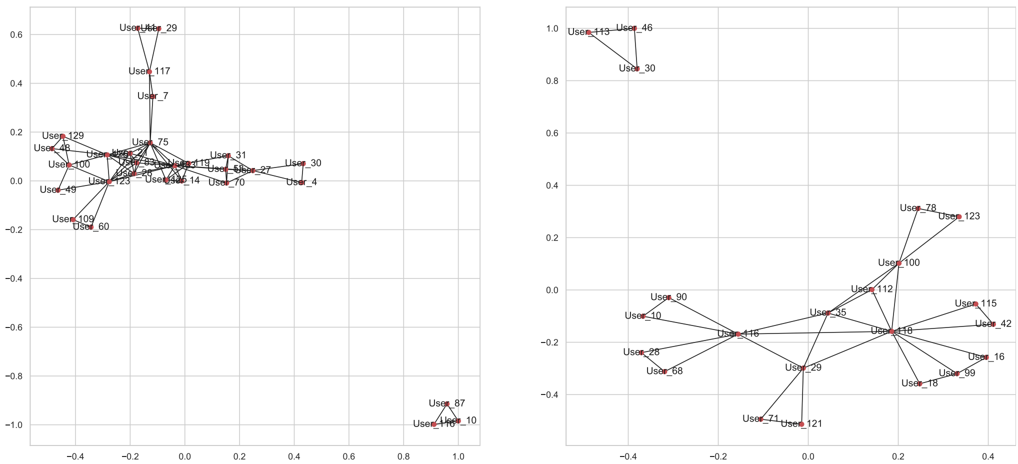

Figure 2: Social network graphs for top-10 (left) and bottom-10 (right) performing teams

5.2 Interview study

We performed semi-structured interviews with professionals from industry to evaluate the usefulness of the method in business settings. The interviewees were approached through email or LinkedIn. We sought for at least medium level knowledge of one of the related fields, process mining (PMin), performance measurement (PFM), and SNA. The information about the interviewees is summarized in Table 6. All interviewees were from EU countries. The interviews took on average 50 minutes. Before the interviews, we sent the description of the method and its example application on the loan management process. During the interviews, we first provided a brief explanation of the method and asked questions on the experience and background of the interviewees. We then proceeded with questions on the usefulness, benefits, opportunities, limitations, and risks. To make sure the interviewees are clear about the goal and application of the method, based on their experiences, we asked them to reflect on the possible applications of the method in their organizations and explained the method further if needed based on their responses. We coded all interview transcripts and performed a thematic analysis to identify common topics and discover unexpected themes (Braun and Clarke, 2006). We summarize the interview results in five topics below based on the identified themes.

Benefits: The interviewees saw potential benefits such as the value of transparent data to make balanced resource allocations among tasks (P1 and P2) and uncovering reasons of performance deviations in different cases (P3 and P4). Increased productivity for the organization by discovering improvement opportunities in team compositions was another expected benefit (P2, P4, and P5).

No	Experience	Current Position	Domain	PMin	PFM	SNA
P1	>20 years	Director	Business process management	●	●	○
P2	>20 years	C-founder	Digitalization, 3D print	◑	◑	◑
P3	>15 years	Head of data science	Anonymous	●	●	○
P4	>10 years	Freelance consultant	Business process management	●	●	●
P5	>20 years	Accountant	Banking	○	◑	○
P6	>10 years	Managing director	Process mining, data science	●	●	◑

Table 6: Information on interviewees and their level of knowledge on related fields

Limitations: All interviewees expressed concerns about legal and privacy issues. They argued on possible constraints of using individual data based on the laws in European Union. Anonymization was suggested not to make personal inferences from the outcomes (P1, P2, and P4). P1 and P4 saw risks in the possible misuse of outcomes to dismiss or assign employees to low-profile tasks. P2 and P3 highlighted the importance of data quality to eliminate wrong or biased outcomes.
Use of the method: The expected uses were understanding deviations among teams and improving poor-performing teams (P1, P4, and P5), improvement of overall team efficiency, focusing on specific metrics to reveal different dimensions of performance (P3), and using results as input for a digital twin of the teams and departments.
Metrics: The use of different types of metrics was considered essential in the method by all interviewees. The inclusion of soft metrics such as employee satisfaction (P2 and P3), contextual metrics such as complexity (P5 and P6), and communication-related metrics (P6) were further suggested.
General Opinion: All interviewees found helpful that the method provides a transparent mechanism to evaluate team performances and help to develop actions. They found the method valuable also because they see a growing demand for process mining. Two interviewees emphasized that it must be used as a supportive tool for decision-making. P1 further expressed the need for explicit recommendations and actions to resolve identified issues. All interviewees indicated concerns about privacy issues, expressing the need to clarify to the employees that the method provides insights at a team level with the purpose of adjusting team compositions rather than personal assessment.

6 Conclusion and discussion

Measuring performance in an objective and accurate way is essential to establish equality among individuals and teams in organizations, improve job satisfaction, and achieve a better working environment. In this study, we aim at contributing to this by developing a performance benchmarking method for teams in organizations based on objective process data. We demonstrated the application of our five-step method and the insights that can be derived on a real-life event log of a loan management process. The usefulness and limitations of the method are further evaluated through interviews. As suggested by the results, our method may enable organizations to benchmark the performances of teams executing business processes and identify improvement areas for the teams. Moreover, the method helps to understand the reasons behind performance differences and take actions to form teams that can perform better in future business process executions. To guide the organizations in objective and accurate team performance evaluation, the method lays down the decisions to be made for data selection and analysis based on the process context.
Implications for research: This study aims to answer the call in previous research on how to design performance measures to encourage cooperation among teams and prevent conflicts between different measures (Crawford and Cox, 1990; Neely, Gregory, and Platts, 2005). By measuring the performance of teams, cooperation moves into the foreground, since the method does not incentivize individuals standing out or outperforming the others. Various dimensions of performance can be taken into consideration through the use of multiple metrics. In this way, this study extends the literature on performance measurement, process mining, and social network analysis (e.g., (Aalst and Song, 2004; Z. Huang, Lu, and Duan, 2012; Pika et al., 2017)). It shows how these research areas can be combined to measure and evaluate the performance of team of employees engaged in the same process.

Implications for practice: The insights obtained through the application of the method can be used to form efficient teams, reward high-performing teams, encourage and guide low-performing teams, and provide objective insights to resolve team-related performance issues. Shapiro (1977) argued that one reason for conflicts among departments is the use of inadequate and contractionary evaluation criteria. Our method can overcome such conflicts by using a range of metrics. The method can be used to improve existing methods on work assignment to resources in process-aware information systems by considering the most efficient team compositions (Eck et al., 2017).

Organizations are increasingly worried about and paying attention to the privacy perceptions of their employees, as confirmed by our interview findings. To overcome the privacy concern, it is important that organizations are transparent to their employees about which data is used and for what purposes (European Group on Ethics in Science and New Technologies, 2018). Our method provides insights on team performance and does not include personal assessment, while increasing transparency and objectivity of performance evaluation. Thus, organizations can overcome the privacy concerns of their employees about the method if they communicate the method well and put in place a governance system to manage these concerns (Reinkemeyer, 2020).

Limitations: Our first limitation is that the method was applied on one dataset. It constitutes an example of how the method can be applied and results can be interpreted. However, since we did not work with the organization providing this dataset and we performed on our assumptions rather than actual process knowledge. If the method was applied by the process experts in the organization, decisions such as metric choices and DEA configuration could be made differently, which would result in different outcomes. Future work should include further evaluation of the method in an organization for validity and utility, in conformance with the design science research methodology (Gregor and Hevner, 2013). Second, in our application, there was low number of cases that could be used after data selection. This is due to the existence of many different resources and composition of them into a high number of teams. There can be much fewer number of resources, and accordingly teams, in some business processes (e.g., Schonig et al. (2018)). In that situation, a high number of cases can be retained in the event log to apply the method, which would improve the reliability. Existence of a high number of resources in a process may also point to issues about team formation. Third, the method considers only activities that are performed on information systems and recorded in event logs. If individuals also perform manual activities while performing processes, relying on this method would not be sufficient for accurate performance evaluation.

Future work: In our method, teams are identified as individuals working on one case together and extracted automatically from the event log. There is a potential to complement the method with other approaches to identify the teams and improve team compositions from different perspectives. In some process contexts, employees working on a process might not be connected with each other as a team and might be only loosely related, for example, if separate organizational units are performing parts of the process. If so, an organization can consider to split a process and the event log accordingly to apply the method on the individuals that are working in one part of the process. Sub-processes defined in the process models or parts of the process that are handled consecutively by the same unit can be used to automatically extract such teams. Another approach is to consider characteristics of individuals to identify teams, such as roles or specific personal capabilities, as in Schonig et al. (2018). Handovers between two activities in a process are points of intense collaboration, and it is important to perform them efficiently for overall process performance (Leyer, Iren, and Aysolmaz, 2020). Another complementary approach can be the identification of teams based on handovers (using the methods in Kumar, Dijkman, and Song (2013a) and Leyer, Iren, and Aysolmaz (2020)) and suggesting improved team compositions considering handover characteristics of teams.

Another future research direction is on conducting wider analysis to unveil the reasons of differing team performance. One potential reason that deserves investigation is the synergy effect, which suggests the performance of a team is more than the estimated total performance of its individual members. Based on the identified reasons, predictive suggestions can be automatically produced for establishing the most effective teams at the start of a process and implemented in process-aware information systems.

References

Aalst, W. M. P. van der (2016). *Process Mining: Data Science in Action*. Springer-Verlag Berlin Heidelberg. ISBN: 978-3-662-49851-4.

Aalst, W. M. P. van der, H. A. Reijers, and M. Song (2005). "Discovering social networks from event logs." *Computer Supported Cooperative Work* 14 (6), 549–593.

Aalst, W. M. P. van der and M. Song (2004). "Mining Social Networks: Uncovering Interaction Patterns in Business Processes." In: *Lecture Notes in Computer Science*, pp. 244–260.

Aysolmaz, B., N. Dau, and D. Iren (2020). "Preventing Algorithmic Bias in the Development of Algorithmic Decision-Making Systems: A Delphi Study." In: *Proceedings of the 53rd HICSS*. Vol. 3, pp. 5267–5276. ISBN: 9780998133133.

Aysolmaz, B., D. M. M. Schunselaar, H. A. Reijers, and A. Yaldiz (2019). "Selecting a process variant modeling approach: guidelines and application." *Software & Systems Modeling* 18 (2), 1155–1178. ISSN: 1619-1374.

Bayardo, R. J., Y. Ma, and R. Srikant (2007). "Scaling Up All Pairs Similarity Search." In: *Proceedings of the 16th International Conference on World Wide Web*. WWW '07. Banff, Alberta, Canada: ACM, pp. 131–140. ISBN: 978-1-59593-654-7.

Behn, R. D. (Sept. 2003). "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63 (5), 586–606. ISSN: 0033-3352.

Bogetoft, P. and L. Otto (2011). *Benchmarking with DEA, SFA, and R*. Vol. 157. New York, NY: Springer New York. ISBN: 978-1-4419-7960-5.

Brass, D. J. (Sept. 1981). "Structural Relationships, Job Characteristics, and Worker Satisfaction and Performance." *Administrative Science Quarterly* 26 (3), 331–348.

Braun, V. and V. Clarke (Jan. 2006). "Using thematic analysis in psychology." *Qualitative Research in Psychology* 3 (2), 77–101. ISSN: 1478-0887.

Burger, A. and J. Moormann (2008). "Productivity in banks: Myths & truths of the Cost Income Ratio." *Banks and Bank Systems* 3 (4), 85–94.

Carrington, P. J., J. Scott, and S. Wasserman (2005). *Models and Methods in Social Network Analysis (Structural Analysis in the Social Sciences)*. Ed. by P. J. Carrington, J. Scott, and S. Wasserman. 1st. Cambridge: Cambridge University Press. ISBN: 9780511811395.

Cooper, W. W., L. M. Seiford, and J. Zhu (2011). *Handbook on Data Envelopment Analysis*. 2nd Edition. Vol. 164. International Series in Operations Research & Management Science. Boston, MA: Springer US. ISBN: 978-1-4419-6150-1.

Crawford, K. M. and J. F. Cox (1990). "Designing Performance Measurement Systems for just-in-time Operations." *International Journal of Production Research* 28 (11), 2025–2036.

Delcoucq, L., F. Lecron, P. Fortemps, and W. M. van der Aalst (2020). "Resource-Centric Process Mining: Clustering Using Local Process Models Resource." In: *SAC '20: Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pp. 45–52. ISBN: 9781450368667.

Dugelova, M. and M. Strenitzerova (2015). "The Using of Data Envelopment Analysis in Human Resource Controlling." *Procedia Economics and Finance* 26, 468–475. ISSN: 22125671.

Dumas, M., M. La Rosa, J. Mendling, and H. A. Reijers (2018). *Fundamentals of business process management*. Springer. ISBN: 978-3-662-56509-4.

Eck, M. L. van, M. Firat, W. P. M. Nuijten, N. Sidorova, and W. M. P. van der Aalst (Oct. 31, 2017). "Human Performance-Aware Scheduling and Routing of a Multi-Skilled Workforce." *Complex Systems Informatics and Modeling Quarterly* (12), 1–21. ISSN: 22559922.

European Group on Ethics in Science and New Technologies (2018). *Future of Work , Future of Society*. Tech. rep. European Commission.

Franceschini, F., M. Galetto, and D. Maisano (2007). *Management by Measurement: Designing key indicators and performance measurement systems*. Vol. 1. Berlin: Springer Berlin Heidelberg. ISBN: 9783540732112.

Franco-Santos, M., L. Lucianetti, and M. Bourne (2012). "Contemporary performance measurement systems: A review of their consequences and a framework for research." *Management Accounting Research* 23 (2), 79–119.

Furht, B. (2010). *Handbook of Social Network Technologies and Applications*. Boston, MA: Springer US. ISBN: 978-1-4419-7141-8.

Gregor, S. and A. R. Hevner (2013). "Positioning and Presenting Design Science Research for Maximum Impact." *MIS Quarterly* 37 (2), 337–355. ISSN: 02767783. DOI: 10.2753/MIS0742-1222240302.

Hanachi, C., W. Gaaloul, and R. Mondi (2012). "Performative-Based Mining of Workflow Organizational Structures." In: *E-Commerce and Web Technologies. EC-Web 2012*. Ed. by C. Huemer and P. Lops. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 63–75. ISBN: 978-3-642-32273-0.

Huang, Z., X. Lu, and H. Duan (June 2012). "Resource behavior measure and application in business process management." *Expert Systems with Applications* 39 (7), 6458–6468.

Islam, S. M. N. and A. Manzoni (2009). *Performance Measurement in Corporate Governance*. Contributions to Management Science. Heidelberg: Physica-Verlag HD. ISBN: 978-3-7908-2169-7.

Koch-Rogge, M., G. Westermann, C. Wilbert, and R. Willis (2014). "How are they doing? The Application of Data Envelopment Analysis to Develop a Strategic Performance Appraisal on Individual Level." In: *A Focused Issue on Building New Competences in Dynamic Environments (Research in Competence-Based Management*. 7th Edition. Emerald Group Publishing Limited, pp. 193–218.

Kratzer, J., R. T. A. Leenders, and J. M. Van Engelen (2010). "The social network among engineering design teams and their creativity: A case study among teams in two product development programs." *International Journal of Project Management* 28 (5), 428–436.

Kumar, A., R. Dijkman, and M. Song (2013a). "Optimal Resource Assignment in Workflows for Maximizing Cooperation." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 8094 LNCS, pp. 235–250. ISBN: 9783642401756.

— (2013b). "Optimal resource assignment in workflows for maximizing cooperation." In: *Business Process Management*. Springer, pp. 235–250.

Levy, P. E. and J. R. Williams (2004). "The social context of performance appraisal: A review and framework for the future." *Journal of Management* 30 (6), 881–905.

Leyer, M., D. Iren, and B. Aysolmaz (2020). "Identification and analysis of handovers in organisations using process model repositories." *2Business Process Management Journal* In Press.

Lin, N. (2001). *Social Capital A Theory of Social Structure and Action*. Cambridge: Cambridge University Press. ISBN: 9780511815447.

Manski, C. F. (Feb. 18, 2000). "Economics Analysis of Social Interactions." *The Journal of Economic Perspectives* 14 (3), 115–136.

Molm, L. D. (Sept. 1994). "Dependence and Risk: Transforming the Structure of Social Exchange." *Social Psychology Quarterly* 57 (3), 163. ISSN: 01902725.

Nakatumba, J. and W. M. P. van der Aalst (2010). "Analyzing Resource Behavior Using Process Mining." In: *Lecture Notes in Business Information Processing*. Vol. 43 LNBIP, pp. 69–80. ISBN: 3642121853.

Neely, A., M. Gregory, and K. Platts (Dec. 2005). "Performance measurement system design: A literature review and research agenda." *International Journal of Operations & Production Management* 25 (12), 1228–1263. ISSN: 0144-3577.

Newell, S., C. Tansley, and J. Huang (2004). "Social capital and knowledge integration in an ERP project team: the importance of bridging and bonding." *British journal of management* 15 (S1), S43–S57.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee (2007). "A design science research methodology for information systems research." *Journal of Management Information Systems* 24 (3), 45–77. ISSN: 07421222. DOI: 10.2753/MIS0742-1222240302.

Pika, A., M. Leyer, M. T. Wynn, C. J. Fidge, T. A. H. M. Hofstede, and W. M. P. van der Aalst (Mar. 23, 2017). "Mining Resource Profiles from Event Logs." *ACM Transactions on Management Information Systems* 8 (1), 1–30.

Plotnikova, V., M. Dumas, and F. Milani (2020). “Adaptations of data mining methodologies: a systematic literature review.” *PeerJ Comput. Sci.* 6 (267). DOI: 10.7717/peerj-cs.267.

Provost, F. and T. Fawcett (2013). *Data Mining for Business*. Ed. by M. Loukides and M. Blanchette. 1st Edition. Sebastopol, CA: O’Reilly Media, Inc. ISBN: 978-1-449-36132-7.

Reinkemeyer, L. (2020). “Challenges, Pitfalls, and Failures.” In: *Process Mining in Action: Principles, Use Cases and Outlook*. Ed. by L. Reinkemeyer. Cham: Springer International Publishing, pp. 37–40. ISBN: 978-3-030-40172-6.

Rodrigues, A., C. Almeida, D. Saraiva, F. Moreira, G. Spyrides, G. Varela, G. Krieger, I. Peres, L. Dantas, M. Lana, et al. (2017). “Stairway to value: mining a loan application process.” *BPI Challenge*.

Saltz, J., N. Hotz, D. Wild, and K. Stirling (2018). “Exploring project management methodologies used within data science teams.” In: *Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018*, pp. 1–5. ISBN: 9780996683166.

Schonig, S., C. Cabanillas, C. Di Ciccio, S. Jablonski, and J. Mendling (May 2018). “Mining team compositions for collaborative work in business processes.” *Software & Systems Modeling* 17 (2), 675–693. ISSN: 1619-1366.

Schönig, S., C. Cabanillas, S. Jablonski, and J. Mendling (2015). “Mining the Organisational Perspective in Agile Business Processes.” In: *Enterprise, Business-Process and Information Systems Modeling*. Ed. by K. Gaaloul, R. Schmidt, S. Nurcan, S. Guerreiro, and Q. Ma. Cham: Springer International Publishing, pp. 37–52. ISBN: 978-3-319-19237-6.

Shapiro, B. P. (Sept. 1977). “Can Marketing and Manufacturing Coexist?” *Harvard Business Review*, 104–114.

Song, M. and W. M. P. van der Aalst (2008). “Towards comprehensive support for organizational mining.” *Decision Support Systems* 46 (1), 300–317. ISSN: 0167-9236. DOI: 10.1016/j.dss.2008.07.002.

Sparrowe, R. T., R. C. Liden, S. J. Wayne, and M. L. Kraimer (Apr. 1, 2001). “Social Networks and the Performance of Individuals and Groups.” *Academy of Management Journal* 44 (2), 316–325. ISSN: 0001-4273.

Swennen, M., G. Janssenswillen, M. J. Jans, B. Depaire, and K. Vanhoof (2015). “Capturing Process Behavior with Log-Based Process Metrics.” In: ed. by P. Ceravolo and S. Rinderle-Ma, pp. 141–144.

Thevendran, V. and M. Mawdesley (Feb. 2004). “Perception of human risk factors in construction projects: an exploratory study.” *International Journal of Project Management* 22 (2), 131–137. ISSN: 02637863.

Van Dongen, B.F. (2017). *BPI Challenge 2017*. en. URL: https://data.4tu.nl/repository/uuid:5f3067df-f10b-45da-b98b-86ae4c7a310b.

Wirth, R. and J. Hipp (2000). “CRISP-DM: Towards a standard process model for data mining.” In: *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*. 24959, pp. 29–39. DOI: 10.1.1.198.5133.

Yang, H. L. and J. H. Tang (2004). “Team structure and team performance in IS development: A social network perspective.” *Information and Management* 41 (3), 335–349. ISSN: 03787206. DOI: 10.1016/S0378-7206(03)00078-8.

Zhao, W. and X. Zhao (2014). “Process Mining from the Organizational Perspective.” In: *Foundations of Intelligent Systems*. Ed. by Z. Wen and T. Li. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 701–708. ISBN: 978-3-642-54924-3.