Towards a Data-driven Framework for Measuring Process Performance

Isabella Kis, Stefan Bachhofner, Claudio Di Ciccio, and Jan Mendling

Vienna University of Economics and Business, Austria
isabella.kis93@gmail.com, s.bachhofner@me.com,
{claudio.di.ciccio,jan.mendling}@wu.ac.at

Abstract. Studies have shown that the focus of Business Process Management (BPM) mainly lies on process discovery and process implementation & execution. In contrast, process analysis, i.e., the measurement of process performance, has been mostly neglected in the field of process science so far. However, in order to be viable in the long run, a process' performance has to be made evaluable. To enable this kind of analysis, the suggested approach in this idea paper builds upon the well-established notion of devil's quadrangle. The quadrangle depicts the process performance according to four dimensions (time, cost, quality and flexibility), thus allowing for a meaningful assessment of the process. In the course of this paper, a framework for the measurement of each dimension is proposed, based on the analysis of process execution data. A trailing example is provided that reflects the expressed concepts on a tangible realistic scenario.

Keywords: Business processes, process analytics, devil's quadrangle

1 Introduction

According to a survey conducted by Müller in 2010, a majority of the questioned companies saw a direct correlation between Business Process Management (BPM) and corporate success [13]. To be able to conduct BPM successfully the performance of a process needs to be measured. Nonetheless, studies have shown that business process analysis has long been neglected in the field of BPM [16,19], as BPM devoted most of the research endeavours on the aspects of process discovery, and process implementation and execution. To be able to analyse a process properly, process performance has to be measured first. A well-established paradigm in that sense is dictated by the so-called devil's quadrangle [4,8]. It shows process performance based on four dimensions: time, cost, quality and flexibility. Those four factors for performance measurement influence each other in a way that it is not possible to improve performance of one dimension without affecting other dimensions, either positively or negatively [4]. An advantage of the approach of measuring process performance with the devil's quadrangle is the possibility to compare changes in the performance over time as visually depicted in Fig. 1.

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

So far, process analysis has been a neglected field of process science and only a few suggestions have been made on how process performance could be measured according to those four dimensions, such as in the case of [8,11,10,20]. Furthermore, to the best of our knowledge, no metrics have been proposed for those dimensions that can be automatically measured over the log data of Business Process Management Systems (BPMSs). However, in the long run process analysis will be crucial for corporate success, thus demanding a framework that allows for a meaningful assessment of a process.

With a focus on the service sector, we propose a framework that suggests how metrics for the devil's quadrangle's dimensions can be derived by using log data generated by a process engine. Our final aim is to help the team involved in a BPM initiative to make informed decisions on the changes to apply to the processes under analysis, driven by factual knowledge stemming from real data. To increase the applicability of the suggested framework, we propose measurements based on values that are most commonly recorded by BPMSs, such as the time and resource allocation of activity executions & incident handling. In the spirit of the idea paper, we focus on the rationale behind the proposed metrics and exemplifications thereof, paving the path for formal and technical treatises. The presented framework is based upon the results of a dedicated investigation on the matter, conducted in the context of a research project in collaboration with PHACTUM Softwareentwicklung GmbH.

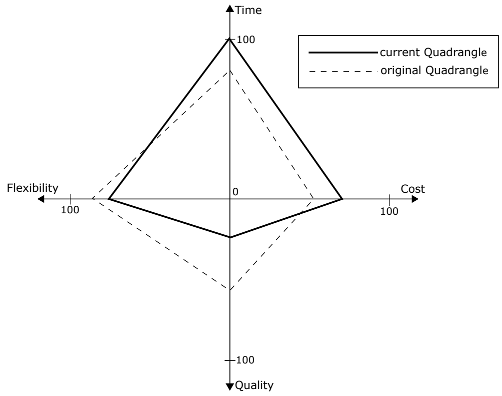

Fig. 1: Changes in process performance [4]

The remainder of the paper is structured as follows: Section 2 proposes a trailing example process and draws preliminary considerations on the analysis; Section 3 provides a framework on how the four dimensions of the devil's quadrangle can be measured by using log data generated by a process engine. Finally, Section 4 concludes the paper and draws some remarks for future research in the field.

2 Preliminaries

Figure 2 depicts an insurance claim process. The process starts with a claim that is received and forwarded (activity A) to a specialist by the secretary of the insurance company. The specialist then assesses the damage (activity B) and writes a damage report (activity C). Subsequently, it has to be decided whether

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

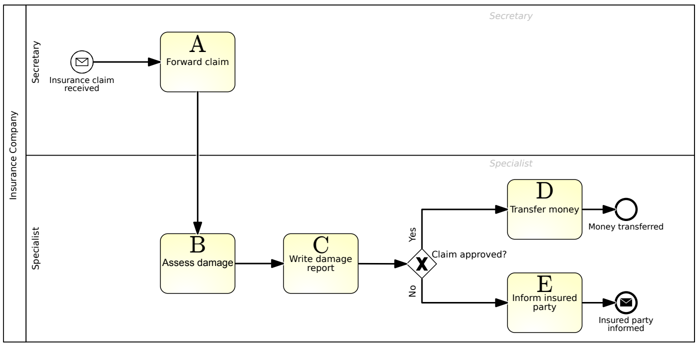

Fig. 2: Example process of an insurance claim

Table 1: Average duration of the insurance claim process

	A	B	C	D	E	Wait. time	Total
Avg. Duration	10	300	150	20	20	460	940

the claim is approved or not. In case of approval, the money to cover the damage is transferred to the policyholder (activity D). If the claim is rejected, the insured party is informed (activity E).

As it can be seen in Table 1, the average duration of the process is 940 minutes, which is equivalent to 15.66 hours. The total average duration consists of the average duration of every single activity (except for activities D and E who are counted as one, as one instance can only take one path, plus the time an instance had to wait for further processing, i.e., the wait time). To sum it up, the length of the observation period roughly corresponds to two working days, assuming that one working day amounts to 7.5 hours. The reference period is by default also set to two working days, i.e. 940 minutes.

The example process and its log data will be exploited in the remainder of the paper to exemplify how the suggested metrics are measured. Before the derivation of the metrics for each dimension though, we draw some preliminary considerations about *(i)* the time span into which the process performance is assessed, and *(ii)* the comparability of the measurements.

For what the first point is concerned, we are interested in the notions of observation time and reference time. The observation time in terms of duration is equivalent to the lead time of a process, namely the time it takes to handle an entire case. To assess the performance of a process, it is crucial to know for how long data on a process needs to be collected in order to allow for a sound statement. Reference time, on the other hand, relates to the past performance of a process or, more precisely, to the period of time for which former process performance is observed. The setting of a reference time enables a comparison of current process performance and past process performance, thus making it possible to further enhance performance assessment. Therefore, the reference

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

period is used to compare the metrics measured during the observation period with past process performance, so that conclusions about the development of the process can be drawn.

To derive benefit from the devil's quadrangle the observation period has to be chosen carefully. This is because it is highly unlikely that a reasonable conclusion about process performance can be drawn from the quadrangle if the observation period is longer or shorter than the actual process time. Following the cycle-time concept from Kanban literature [3], the average total duration of the process is used as a basis of calculation – consequently, data on the process has to be collected first. To ensure that the amount of data collected is sufficient, we recommend that the process owners be consulted. They would know how long the process lasts on average and can recommend an adequate period of time for data collection. Once the average total duration of the process has been determined, a safety margin in the form of the standard deviation will be added. Of course the time frame for the observation period can be modified by the user. The observation period calculated by the system is merely a default setting and has to be seen as a recommendation for the user. To calculate the duration of the observation period for the example process, the average total duration of the process has to be determined. According to Table 1 the average observation period should be set to 940 minutes. Please notice that for reasons of simplicity no safety margin was added.

In Fig. 1 the reference period is represented by a dashed line. The reference period depicts the past performance of a process for a predefined period of time. By default the reference period comprises the same time frame as the observation period. This setting can again be changed by the user according to the current evaluation needs. The reference period for the example process is equal to the observation period's duration and amounts for 940 minutes. Both the current quadrangle (i.e. observation period) and the original quadrangle (i.e. reference period) are put on top of each other in order to make comparison possible.

To guarantee the comparability of the devil's quadrangle's four dimensions their values should move within the same scale. In this paper it has been decided upon a scale ranging from 0 to 100 percent. The more the value of one dimension approaches 0, the worse the respective dimension performed.

3 Approach

Throughout this section, we define the metrics that we associate to each dimension of the devil's quadrangle: Section 3.1 deals with the time dimension, Section 3.2 is concerned with the cost dimension, Section 3.3 focuses on the quality dimension, and Section 3.4 discusses the flexibility dimension.

3.1 Time Dimension

When measuring the time dimension of a process, we are interested in how much time is dedicated to the carry-out of the tasks of a process instance. Consequently, we focus on the service time, i.e., the time the resources spend on

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

Table 2: Measurement of the time and cost dimensions

	Activity					Wait. time	Lead. time	Service time	Service/time ratio
	A	B	C	D	E				
Run 1	10	240	60	20		360	690	330	47.83%
Run 2	5	300	120		20	240	685	445	64.96%
Run 3	20	360	120	10		240	750	510	68.00%
Run 4	10	240	120	10		180	560	380	67.86%
Run 5	10	180	60		20	360	630	270	42.86%
Run 6	25	300	240		10	360	935	575	61.50%
Run 7	10	300	120	30		300	760	460	60.51%
Run 8	10	240	60	10		60	380	320	84.21%
Run 9	5	360	180	20		360	925	565	61.08%
Run 10	5	360	180		10	360	915	555	60.66%
Total Runtime	110	2.880	1.260	100	60				

actually handling a case [8]. The service time of a case is then compared with the case's lead time, thus indicating how much of the total time an instance takes to finish is spent on actual work. The higher the service time ratio, the better, as more time has been spent on actually handling a case and less time was lost due to a process instance being at a resting stage.

As it is very likely that more than one case is examined during the observation period, we once again calculate the service time ratio for each process instance, i.e., the comparison of a case's service time with its respective lead time, and calculate a median for all the single values. The resulting median is then transferred to the time axis of the quadrangle.

In order to generate data for the calculations regarding the insurance claim process, we ran an example of ten instances of a process (see Table 2). First, the lead time of the process, consisting of the duration of each activity and the wait time, i.e., the time an instance waited for further processing, was calculated. After that, the service time (i.e. the time a process instance was actually handled) was computed. These two steps were taken in order to be able to gather the service time ratio, which indicates how much time of the process was spent on actual work. In the end, the service time ratios for each run were sorted in descending order to calculate the median for the time dimension. The resulting median for our computation is 61.29%.

3.2 Cost Dimension

To calculate the process costs, personnel expenses for each process task are stored in a variable. Then the expenses for each task of the process are added up to a total value. A justification for the use of personnel cost for calculating a process' cost is seen in the importance of this type of cost for organisations. Personnel expenses normally represent the most relevant cost type of the service sector – in production industry they are the second most important cost type [6].

The personnel costs are most likely stored in a central database, which contains the salary of every employee. Through the integration of such an information in the database with the data logs of the BPMS, the current hourly payment of each employee can be calculated. It is important to notice that for cost calculation the actual personnel costs have to be used. The term *actual personnel costs* refers to direct payments to employees increased by continued payment of salaries (in case of holiday, sick leave and bank holidays), holiday pay, Christmas bonuses, the employer's social security contributions, overtime rates and other personnel costs [6].

As the event log stores information about how long an employee has been working on a task, a viable measurement of the process' cost can be achieved by multiplying the actual labour costs per hour by the processing time per task. The costs will then be determined for the selected observation period. The result is then compared with the total costs of the organisation for the same observation period resulting in a percentage value that can be transferred to the cost axis of the devil's quadrangle. The higher the value the worse the process performed with regard to the cost dimension (i.e. a high value means high personnel costs compared to total costs). However, it has already been mentioned in Section 2, we want all the axes to have a uniform meaning: The closer the value is to 100% the better process performance is rated. This is why the value received from the previous division has to be inverted before transferring it to the cost axis.

Considering the insurance claim process example, it is assumed that the costs for the secretary amount to \$20 per hour and that the specialist is paid \$40 per hour. Moreover, it is exactly known which activities are handled by whom. With this in mind, the total duration of activity A is multiplied by the hourly costs of the secretary, whereas the duration of the remaining activities is multiplied by the hourly rate of the specialist (the total duration can also be extracted from Table 2). The resulting sum is then compared with the total cost of the company, which we estimated with \$4.000 per working day. The value for the cost dimension thus amounts to the inverted ratio: Its value, 27.42%, can be depicted on the quadrangle.

We remark here that we assume a complete knowledge of the work of resources, with information on the assigned task and duration of the carry-out thereof. This is a reasonable assumption in case a BPMS is supporting the process execution. Otherwise, the conduction of tasks in parallel, or the interruptions during task handling, holidays, weekends, etc., need a substantial amount of effort to be considered [1].

3.3 Quality Dimension

When measuring the quality of a process we want to consider two different aspects: First, we want to examine whether the process finished as planned. We will refer to this first aspect of quality as *outcome quality*, as it can help to judge the path a process instance took to finish the process. Second, it is to be checked whether any incidents (i.e., technical errors that can occur during process execution) were created. We will henceforth refer to it as *technical quality*. After both subdivisions of quality have been evaluated they are combined and transferred to the quality axis of the devil's quadrangle.

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

Table 3: Process traces

Run 1	A	B	C	D
Run 2	A	B	C	E
Run 3	A	B	C	D
Run 4	A	B	C	D
Run 5	A	B	C	E
Run 6	A	B	C	E
Run 7	A	B	C	D
Run 8	A	B	C	D
Run 9	A	B	C	D
Run 10	A	B	C	E

Table 4: Measurement of the technical quality

	Incidents	Elements	Incident rate
Run 1	0	8	0%
Run 2	5	8	62.5%
Run 3	0	8	0%
Run 4	1	8	12.5%
Run 5	1	8	12.5%
Run 6	0	8	0%
Run 7	2	8	25%
Run 8	0	8	0%
Run 9	0	8	0%
Run 10	1	8	12.5%

Outcome Quality. The measurement of the outcome quality serves to assess the course a process instance takes to reach the end of a process. This implies the existence of one or more ideal paths through the process. Yet it would be very time-consuming to assess each process element's affiliation to the ideal path as in most of the times there are different ways through the process an instance can take. Moreover there could be various ideal paths.

Information on the termination of the process instance, typically depicted as end-events in executable process models, should be added. It should include the information whether the achieved outcome was positive or negative. Then, the number of end events that led to a positive outcome of the respective process instance are compared with the total number of process instances executed during the observation period, resulting in a percentage value that can be transferred to the quadrangle's axis. Within the scope of this paper, we assume that a process has at least two end elements, of which one has a positive and the other a negative outcome. The more end events a process has the higher the chance to make a fundamental statement about the process' ideal path(s). Other ways can indeed be adopted to mark the executions as reaching the expected process goal or not. For instance, a process does not necessarily have more than one end event. In this case there could be an exclusive or inclusive split at some point. Depending on the path the process instance takes after that split it is decided whether the decision had a positive or negative impact on the process.

Therefore, the approach to measure the outcome quality should be seen as a starting point for further research. Ideally, it will be possible in the future to identify a path quality, not just assessing the end elements of a process, but rather evaluating whether a specific process element belongs to the ideal path.

Table 3 shows all the paths that the simulated instances took through the insurance claim process. In order to calculate the outcome quality, the meaning for each end point of the process has to be defined. As an insurance company most likely prefers not to pay for a damage claimed by a policy holder, the end event *Money transferred* (subsequent to activity D) has a negative impact on process quality, whereas the end event *Insured party informed* (following activity E) has a positive impact. In Table 3 it can be observed that four out of ten runs ended with a rejection of the insurance claim, which has a positive meaning

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

for the insurance company. At that stage, the number of positive end events is compared with the total number of process instances within the observation period. The four positive end events thus are divided by the total of ten, i.e., the number of process instances in the observation period. The result is an estimated outcome quality of 40%.

Technical Quality. To enable the assessment of a process' technical quality the number of incidents within a predefined observation period can be counted. The more incidents registered for a specific period of time, the worse the process performed in terms of technical quality. In the end, all incidents recorded in the observation period are compared with the total number of elements in a process.

The following example should help to better illustrate this procedure. It is assumed that a process instance records 20 events in the log. During the process execution five incidents are thrown. The technical quality given by the inverted ratio of incidents per process results in a value of 75% for that process instance. Afterwards, a median is calculated for all the values of the separate process instances. The resulting median is then transferred to the quadrangle's quality axis.

Table 4 summarises all the incidents that were registered during the run of one process instance of the example insurance claim process. The number of incidents is then compared with the number of elements that occurred within the same process instance, resulting in a percentage value. The higher this value, the higher the number of incidents within one process instance. To gather a value that represents the technical quality of the whole observation period, the median for all incident rates is calculated. Thus, the technical quality of the process equals 87.5% We remark that the ratio is inverted so as to keep consistency and comparability of the metrics: the more the values on the scale approach 100 %, the better the process performed. In contrast, a higher incident rate means lower technical quality, thus requiring for an inversion of the original result.

Combining Outcome and Technical Quality. To transfer a single value to the respective axis of the quadrangle, we combine the aforementioned quality measurements into a single one. This is achieved by assigning a weight, which can be chosen according to the interest of an organisation, to each of the two quality metrics, i.e., outcome quality and technical quality. Then the values resulting from the measurement of each dimension are multiplied with their respective weights in order to compute a single value that can be transferred to the quality axis of the devil's quadrangle.

For the example process, it was chosen to weigh both outcome and technical quality with 50% which results in a combined value of 66.88%, which can be transferred to the quadrangle's quality axis.

3.4 Flexibility Dimension

According to the Cambridge Dictionary, flexibility generally is "the quality of being able to change or be changed easily according to the situation". In BPM-related literature different specific ways of how to define flexibility can be found [9,12,5]. Accordingly, we adopt various definitions of flexibility, each contributing to a different aspect of the considered process.

In particular, we build upon the notion of run time flexibility as defined in [8]. Run time flexibility is the ability to react to changes while a workflow is executed. We identify two main components that contribute to it. We first focus on the concept of volume flexibility, namely the ability to handle changing volumes of input, rephrasing the definition of [9]. The paper discusses, among other things, a framework for IT-flexibility which can be divided into three dimensions. The first dimension, which is called "Flexibility in Functionality", is concerned with the reaction of a system to changing input conditions. The system is considered flexible if it can withstand varying input conditions. According to [17] flexibility is the maintenance of a stable structure in the face of change, where the structure is intended as what stands between the input and output. In the light of the above, we define the flexibility as follows: *Flexibility is the ability to keep the processing speed of the single instances at an approximately constant level even though the workload (i.e., the input) has increased (or decreased) significantly.* Even though our definition of flexibility slightly differs from the one of volume flexibility in [8] we will henceforth use this term to refer to the ability of a process to keep the instance's processing speed at a constant level when there has been an increase in work.

The other component of the run time flexibility concerns the ability to resolve system exceptions that are thrown during the execution of a process. By incident we mean a technical problem occurred during the BPMS-aided process execution. Such an aspect is of particular relevance in several scenarios where BPMSs are used in practice. This particular kind of flexibility will be thus referred to as *technical flexibility.*

The remainder of this section will be concerned with a more detailed description of the aforementioned components of the run time flexibility. Moreover it will be stated how metrics for the devil's quadrangle can be measured for each of the two.

Volume Flexibility. We define volume flexibility as the ability to guarantee a constant handling of process instances if there has been a change in the workload. To facilitate the understanding of this flexibility concept, we consider the case of the insurance company. In times of natural disasters, the number of insurance claims would increase significantly, resulting in a higher workload as well [2]. If the insurance company manages to adapt to the changed conditions it is considered flexible.

The measurement of volume flexibility is based upon the lead time of a process. Flexibility is examined from a holistic perspective here, which is highly suitable for the assessment of process performance at a glance. Within the scope of this paper the existence of a BPMS is assumed. Therefore each process instance is assigned an ID which allows for a proper estimation of when the process instance started and finished. Consequently, this knowledge enables us to make an exact statement of the process instance's lead time. From the measurements

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

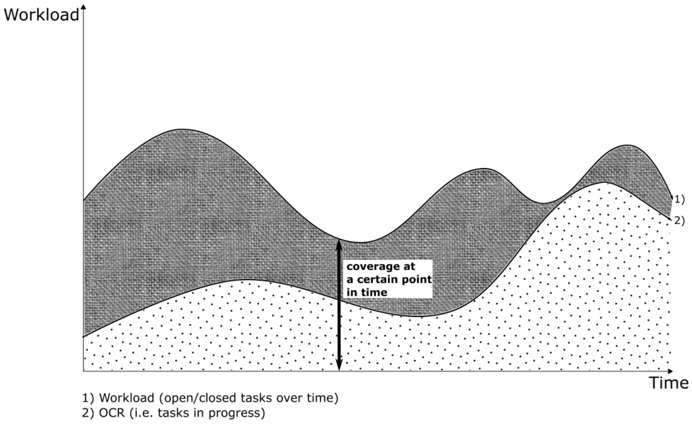

Fig. 3: Measurement of the flexibility metrics

taken in the initial phase of the process analysis we know the planned average lead time of the process.

A first approach to the concept of flexibility is the calculation of an open-closed-ratio (OCR) for a previously defined observation period. This ratio is computed by comparing cases in progress (open cases) with completed cases (closed cases). When there has been an increase in the workload and the values for open and closed cases balance roughly, it can be presumed that there is a constant handling of cases. If, on the other hand, there is no balance between the two values, it can be concluded that the process lacks flexibility.

The only problem of measuring the flexibility dimension as suggested above is that the results could be corrupted. This is owing to the fact that the workload level is not taken into account. To put it differently, what would happen if the workload does not change, hence remains at a constant level, and the OCR indicates a constant handling of cases? It could be assumed that the organisation is highly flexible even though the workload remained stable. However, this does not correspond to the definition of flexibility, rooted in the ability to adapt to a changed or new situation. For this reason an additional factor has to be included in the measurement of flexibility: the workload itself.

As it can be seen in Fig. 3, two factors are taken into consideration for the measurement of process flexibility: *(i)* the OCR, and *(ii)* the workload, i.e, the cases in progress. The OCR shows the open and closed cases of an observation period, i.e., the changing workload over time. If the workload increases (resp. decreases) over time, the OCR has to rise (fall) too in order to be able to speak of a highly flexible process. In contrast, if the OCR remains on the same level this is a sign of lacking flexibility, because the process is apparently not able to adapt to changing conditions.

To receive a percentage value for the volume flexibility, the coverages both of the workload and the OCR curve have to be compared. The higher the coverage,

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

the more flexible the process is, as this indicates that the OCR is able to adapt to the changing workload conditions. However, we remark that the OCR will not rise immediately after an increase in the workload, because the cases take a certain time to finish – at least the average lead time. It still has to be considered that both the workload curve and the curve representing the OCR could exactly coincide, even though the number of open and closed cases over time did not increase recently, i.e., the backlog remained on the same level. Again we are confronted with a case where an organisation faces steady workload, which does not correspond to our definition of flexibility. Our understanding of flexibility is that an organisation is able to adapt to changing conditions. But where there is no change, there can be no reaction either. We thus integrate a warning signal that indicates an increased (decreased) amount of cases in progress (or congruent areas below the curves) but no significant rise (fall) of the backlog curve. The user should then be enabled to switch to a more detailed view where both curves are shown, as suggested in [14].

We report two examples showing different ways to measure the volume flexibility, both complying with the described rationale yet tackling the computation from two different perspectives: the first one measures volume flexibility in terms of the total duration per case, the second one adopts a more global perspective and focuses on the number of cases that were opened and closed within the observation period. Both examples refer to the example insurance claim process.

For what the first computation strategy is concerned, Table 5 shows the open cases within the observation period of 940 minutes. It is assumed that the usual number of open cases is ten. It can be then recognised that there are five additional cases to handle with respect to the expectations. This implies that the workload has increased and the measurement of the process' flexibility can be started. To consider the process flexible, each case has to finish within the average lead time of the process. In the "Lead time" column the actual lead time of the respective instance is reported. The average lead time, based on our calculation for the framework, is given in column "Target lead time". In another step, target lead time and actual lead time of every instance are compared, showing that in total the 15 instances took 225 minutes longer than planned to finish. Expressed as a percentage, the process took 1.60% longer than initially planned, thus reducing volume flexibility to 98.40%. The second strategy to measure the volume flexibility metric refers to open and closed cases in the observation period. In Table 6 the open cases in the observation period are reported. Under the assumption that the normal number of open cases is ten, it can be recognised that there are additional five cases to handle which signals a higher workload. In order to be considered flexible, all 15 cases have to be closed within the average lead time of the process. However, as can be gathered from Table 6, only five cases were closed, meaning that 66.67% of the cases are still open for processing and thus reducing volume flexibility to 33.33%.

The purpose of the variation in the calculation of volume flexibility for the example process is to show how different viewpoints influence the outcome of metrics measurement. It can be recognised that the results for volume flexibility

Table 5: Measurement of the process flexibility

Open case	Lead time	Target lead time	Difference
No. 1	690	940	-250
No. 2	685	940	-255
No. 3	750	940	-190
No. 4	560	940	-380
No. 5	630	940	-310
No. 6	950	940	+10
No. 7	940	940	0
No. 8	960	940	+20
No. 9	1.000	940	+60
No. 10	960	940	+20
No. 11	1.200	940	+260
No. 12	1.500	940	+560
No. 13	1.300	940	+360
No. 14	1.000	940	+60
No. 15	1.200	940	+260
Total	14.325	14.100	+225

Table 6: Measurement of the volume flexibility

Open cases	Closed cases	Ratio
15	5	33,33%

differ considerably when comparing the two versions. An organisation therefore has to decide how volume flexibility is measured according to the duration of the process or the number of processed cases, depending on the perspective that is to be emphasised.

Technical Flexibility. When it comes to the measurement of technical flexibility, the number of incidents thrown within a process over a predefined observation period has to be observed. As already stated before, incidents are technical errors which can occur during the execution of a process. To measure the technical flexibility we want to find out how long it takes to resolve one incident. In this case, we are interested in the reaction time. The reaction time for resolving an incident (or the sum of all the time intervals spent for each incident, in case more than one occurred within a process instance) is then compared with the lead time of the corresponding process instance. It is thus indicated how much time of the process execution is dedicated to the handling of technical issues. This way various ratios of the reaction time are received. In order to be able to transfer the ratios to the quadrangle's axis, we calculate a median for them. Before transferring the resulting value to the quadrangle, it is inverted. In practical real-world scenarios, the reaction time for resolving an incident is sometimes not accounted within the lead time. In such a case, the computed value would fall below zero, which is detrimental to our representation, because it aims at normalizing every measurement in the 0-100% range. To circumvent this problem, the incident reaction time can be added to the lead time in the computation.

We remark here that in our proposal the metrics for the technical flexibility deal with incidents as well as in the case of the technical quality. Nevertheless, we aim at representing with flexibility a perspective that mostly pertains the area of management within the organisation, whereas quality is intended to be perceived also outside the scope of the process owners [4], hence all the stakeholders. Owing to this, we look at the reaction time to handle incidents as a flexibility indicator, because it is an information mostly kept within the organisation. The time the

Table 7: Measurement of the technical flexibility

	Lead time	Incidents	Total reaction time	Reaction time ratio
Run 1	690	0	-	-
Run 2	685	5	60	8.76%
Run 3	750	0	-	-
Run 4	560	1	10	1.79%
Run 5	630	1	30	4.76%
Run 6	935	0	-	-
Run 7	760	2	10	1.32%
Run 8	380	0	-	-
Run 9	925	0	-	-
Run 10	915	1	40	4.37%

delegated team spent on handling incidents is indeed an internal information that is usually not publicly shown. Ideally, the incident handling time is completely transparent to clients and partners. In contrast, we interpret the number of occurred incidents as an indicator that can be reverberated also outside the organisation, because of the possible disruptions caused thereby.

Table 7 depicts the ten instances of the provided example process, with the addition of the incidents thrown during the execution, and the time needed to resolve the incident in minutes. Subsequently a ratio for the reaction time is calculated, which in the end results in a median of 4.76%. After inverting the ratio, the value for technical flexibility is equal to 95.63%.

Combining Volume and Technical Flexibility. So far we described two different metrics for the measurement of flexibility, namely volume and technical flexibility. In order to have a single metric accounting for both, our suggestion is again to assign a weight to each of the two flexibility components. An organisation, for example, may deem as very crucial to resolve incidents as quickly as possible. Therefore it would weigh the technical flexibility with 80% (out of 100%) and the remaining 20% would be assigned to volume flexibility. The values resulting from the measurement of each dimension would then be multiplied by their respective weights and summed up in order to form a single value that can be transferred to the flexibility axis of the devil's quadrangle.

If such a calculation is conducted for the example process with a weight of 50% each, the flexibility value would amount to 97,02%, in case the first measurement strategy for the volume flexibility is adopted, or to 64,49%, in case the second one was used.

4 Conclusion

Within this paper, suggestions for the measurement of metrics for the four dimensions of the devil's quadrangle have been made, exclusively based on the inspection of process data. All suggested calculations are made under the assumption that the data are extracted from a BPMS event log running the process. The log data used for metrics measurement has been chosen under the condition that it can be extracted from almost any BPMS. For the calculation

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

of a value for the time dimension, the activity execution time is considered as reported in the log. The activity execution time is needed again in order to measure the cost metric. The additional piece of information needed in that case are the personnel expenses. Due to the amplitude of interpretations that can be given to the quality and flexibility dimensions in particular, we identified a combination of metrics, each singularly considering different aspects thereof. For the assessment of the outcome quality, a comparison between the positive and the negative terminations of the process instances is compared. For the measurement of the technical quality, the number of incidents within a process instance is considered. The quality dimension is ultimately assessed as a linear combination of the aforementioned ones. Process flexibility is also assessed as a weighted sum of two different components: the volume flexibility and the technical flexibility. The former is measured on the basis of a comparison between open and closed cases during the observation period. The latter is likewise based on the reaction time to incidents.

It is in our plans to extend the suggested framework towards further refinements and possibilities to customise the measurements. For instance, not only personnel costs but also total process costs should be considered for the cost dimensions, for instance by means of the activity-based costing model. For what the quality dimension is concerned, we would also consider alternative criteria beyond the final outcome or the registered technical incidents, e.g., an *enactment quality* based on the number of times exceptional paths were taken, compared to the expected course of the process unfolding. Furthermore, we are investigating how to better include the concept of dynamics in the flexibility dimension. The proposed metrics indeed average the ratio of values in the observation period (closed v. open cases, or incident reaction time v. lead time). However, the flexibility is arguably concerned with the responsiveness to *changes*, hence the suggestion for the analysis of trends. The exploitation of mathematical devices such as derivatives to be applied on the ratios demonstrate suitable and are in fact currently under investigation. In this paper, we described a theoretical framework for the measurement of the suggested metrics. Future work will be particularly concerned with its implementation supplemented by expert interviews, so as to conduct a thorough evaluation of the proposed approach on real-world use cases. From this perspective, the recent work of Nguyen et al. [14] for the staged process performance mining shows promising integration opportunities to automate the information extraction and processing needed by our framework. Moreover, it is in our plans to investigate the integration of existing approaches in literature such as SERVQUAL [15] to refine the definition and measurement of the quality dimension, and the SCOR metrics [18] to further investigate the interplay of internal and company-wide processes. Finally, we remark that due to the advanced globalisation, processes too will become more interconnected [7]. Consequently, there is the need to take process performance measurement to the next level and not only assess one single process but also to recognise the interplay of processes organisation-wide, if not beyond company boundaries.

Pre-print copy of the manuscript published by Springer International Publishing
(available at `link.springer.com`)
identified by DOI: `10.1007/978-3-319-59466-8_1`

References

1. van der Aalst, W.M.: Business process simulation revisited. In: EOMAS. pp. 1–14. Springer (2010)
2. van der Aalst, W.M., Rosemann, M., Dumas, M.: Deadline-based escalation in process-aware information systems. Decision Support Systems 43(2), 492–511 (2007)
3. Dickmann, P.: Schlanker Materialfluss : mit Lean Production, Kanban und Innovationen. Springer Vieweg, Berlin, Heidelberg, 3rd edn. (2015)
4. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management. Springer (2013)
5. Gong, Y., Janssen, M.: Measuring process flexibility and agility. In: ICEGOV. pp. 173–182. ACM (2010)
6. Horsch, J.: Kostenrechnung: Klassische und neue Methoden in der Unternehmenspraxis. Springer-Verlag (2015)
7. Houy, C., Fettke, P., Loos, P., van der Aalst, W.M., Krogstie, J.: BPM-in-the-large – towards a higher level of abstraction in business process management. In: E-Government, E-Services and Global Processes, pp. 233–244. Springer (2010)
8. Jansen-Vullers, M., Loosschilder, M., Kleingeld, P., Reijers, H.: Performance measures to evaluate the impact of best practices. In: BPMDS workshop. vol. 1, pp. 359–368. Tapir Academic Press Trondheim (2007)
9. Knoll, K., Jarvenpaa, S.L.: Information technology alignment or "fit" in highly turbulent environments: the concept of flexibility. In: SIGCPR. pp. 1–14. ACM (1994)
10. Kronz, A.: Managing of process key performance indicators as part of the aris methodology. In: Corporate performance management, pp. 31–44. Springer (2006)
11. Kueng, P.: Process performance measurement system: a tool to support process-based organizations. Total Quality Management 11(1), 67–85 (2000)
12. Kumar, R.L., Stylianou, A.C.: A process model for analyzing and managing flexibility in information systems. European Journal of Information Systems 23(2), 151–184 (2014)
13. Müller, T.: Zukunftsthema Geschäftsprozessmanagement. Tech. rep., PricewaterhouseCoopers AG Wirtschaftsprüfungsgesellschaft (2011)
14. Nguyen, H., Dumas, M., ter Hofstede, A.H., La Rosa, M., Maggi, F.M.: Business process performance mining with staged process flows. In: CAiSE. pp. 167–185. Springer (2016)
15. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: Servqual: A multiple-item scale for measuring consumer perc. Journal of retailing 64(1), 12 (1988)
16. Recker, J., Mendling, J.: The state of the art of business process management research as published in the BPM conference. Business & Information Systems Engineering 58(1), 55–72 (2016)
17. Regev, G., Wegmann, A.: A regulation-based view on business process and supporting system flexibility. In: CAiSE. vol. 5, pp. 91–98. Springer (2005)
18. Stephens, S.: Supply chain operations reference model version 5.0: a new tool to improve supply chain efficiency and achieve best practice. Information Systems Frontiers 3(4), 471–476 (2001)
19. Van Der Aalst, W.M.: Business process management: A comprehensive survey. ISRN Software Engineering 2013 (2013)
20. Venkatraman, N., Ramanujam, V.: Measurement of business performance in strategy research: A comparison of approaches. Academy of management review 11(4), 801–814 (1986)

The final version of the paper is identified by DOI: `10.1007/978-3-319-59466-8_1`

References

Kis, Isabella, Stefan Bachhofner, Claudio Di Ciccio, and Jan Mendling (2017). "Towards a Data-Driven Framework for Measuring Process Performance". In: *BPMDS/EMMSAD*. Ed. by Iris Reinhartz-Berger, Jens Gulden, Selmin Nurcan, Wided Guédria, and Palash Bera. Vol. 287. Lecture Notes in Business Information Processing. Springer, pp. 3–18. ISBN: 978-3-319-59466-8. DOI: `10.1007/978-3-319-59466-8_1`.

BibTeX

```
@InProceedings{   Kis.etal/BPMDS2017:DataDrivenFrameworkforMeasuringProcessPerformance,
  author        = {Kis, Isabella and Bachhofner, Stefan and Di Ciccio,
                  Claudio and Mendling, Jan},
  title         = {Towards a Data-Driven Framework for Measuring Process
                  Performance},
  booktitle     = {BPMDS/EMMSAD},
  year          = {2017},
  pages         = {3--18},
  publisher     = {Springer},
  crossref      = {BPMDS-EMMSAD2017},
  doi           = {10.1007/978-3-319-59466-8_1}
}
@Proceedings{     BPMDS-EMMSAD2017,
  title         = {Enterprise, Business-Process and Information Systems
                  Modeling: 18th International Conference, BPMDS 2017, 22nd
                  International Conference, EMMSAD 2017, Held at CAiSE 2017,
                  Essen, Germany, June 12-13, 2017, Proceedings},
  year          = {2017},
  editor        = {Reinhartz-Berger, Iris and Gulden, Jens and Nurcan, Selmin
                  and Gu{\'e}dria, Wided and Bera, Palash},
  volume        = {287},
  series        = {Lecture Notes in Business Information Processing},
  publisher     = {Springer International Publishing},
  isbn          = {978-3-319-59466-8},
  doi           = {10.1007/978-3-319-59466-8}
}
```