

## Design of a Process Mining Alignment Method for Building Big Data Analytics Capabilities

Lukas Pfahlsberger  
WU Vienna, Austria  
[lukas.pfahlsberger@wu.ac.at](mailto:lukas.pfahlsberger@wu.ac.at)

Jan Mendling  
WU Vienna, Austria  
[jan.mendling@wu.ac.at](mailto:jan.mendling@wu.ac.at)

Andreas Eckhardt  
University of Innsbruck, Austria  
[andreas.eckhardt@uibk.ac.at](mailto:andreas.eckhardt@uibk.ac.at)

### Abstract

*Process mining is a big data analytics technique that supports business process management in an evidence-based way. Nowadays, companies struggle to build the required capabilities that lift process mining beyond technical proof-of-concept implementations. As research on process mining is largely limited to algorithm design and project management recommendations, current research does not understand well how process mining and complementary resources and capabilities can be aligned. By understanding those interrelations, companies learn to leverage their organizational potential during the execution of process mining more effectively and efficiently. In this paper, we address this research gap by using the design science research approach to develop a process mining alignment method. Our method supports companies mapping their individual technical requirements of process mining to their underlying organizational resources. We evaluate our method through a series of interviews with IT consultants.*

### 1. Introduction

Business process management has recently become increasingly evidenced-based thanks to the availability of ubiquitous event logs and corresponding process mining (PM) tools and techniques [1]. PM represents a specific kind of big data analytics (BDA) technique, which addresses companies' needs to continuously rethink existing processes that tend to become obsolete rapidly [2, 3, 4] and to reduce the tremendous efforts of manual documentation, maintenance, and optimization. In this way, PM contributes to a company's ability to flexibly change its organizational processes, which, according to Kim et al. [5], "indicates its readiness to undergo other radical reconfigurations".

Even though the potentials of PM are well-supported through various case studies [6], many companies struggle on their journey of leveraging these potentials

on an organizational scale. This underutilized potential becomes even more critical when considering the fast-growing amount of event-log data available but also the increased pace of process change. Various recommendations have been presented for organizing PM projects [7, 8], but these largely have taken a technical or operational perspective. It has been argued in strategic management research that a key challenge is the bundling and strategic alignment of the relevant resources [9, 10]. This implies that PM can be hardly understood solely as a BDA technique, and that organizations have to systematically develop a dedicated BDA capability (BDAC). This BDAC in turn forms the basis for the actual application of PM in individual projects in an efficient and effective way. The strategical alignment of the BDAC as well as the associated application of PM on end-to-end processes can then provide a sustainable competitive advantage through improved processes [11, 12, 13, 14].

In this paper, we address the research gap of how to strategically align a BDAC with PM by designing an alignment method. Our method supports companies mapping their individual technical requirements of PM to their underlying organizational resources. For this purpose, PM is decomposed into its technical configuration dimensions. Subsequently, this configuration dimensions are mapped to organizational resources. This mapping then highlights what is necessary for an effective and efficient application of PM inside the organization. We conducted our research following the inductive reasoning of design science originated by Hevner et al. [15]. Our PM alignment method was subject to build-and-evaluate iterations. It was built on the step-by-step assessment of PM application field, focal perspective, usage types, online or offline implementation, and different contexts. We evaluated our method based on a series of semi-structured interviews with IT consultants and the application of the empirical evaluation method of Gioia [16].

The rest of the paper is structured as follows. Section

2 discusses the theoretical background focusing on the origin and development of process-oriented BDAC and PM configuration dimensions. Section 3 describes our research design grounded in design science and the complementing qualitative evaluation method. Section 4 reports the actual design of our PM alignment method. Section 5 discusses the contributions and limitations before Section 6 concludes.

## 2. Research background

PM research so far largely focused on operational project management and guidelines for conducting concrete projects [8, 7]. Moreover, companies often struggle to align their initial BPM activities with strategic objectives [17]. We believe that part of the problem is the fact that PM is too little understood as a puzzle piece for developing organizational capabilities. Therefore, we discuss PM as an enabling facet of a process-oriented BDAC. To develop such a capability, various resources and contextual factors have to be taken into account. We examine how the underlying resources and contextual factors in conjunction with different configurations of PM shape and align a BDAC.

### 2.1. Big data analytics capabilities and related organizational resources

The term BDA is composed of two constituent parts – *BD* and *analytics*. *BD* as a standalone term is widely discussed in literature. There are numerous definition approaches because its demarcation can be ambiguous in different contexts. Nevertheless, the prevailing definition is the categorization into various *Vs*, also called dimensions, of *BD*. The most common dimensions are volume, variety, and velocity [18, 19, 20]. This has been extended by the dimensions value [21], veracity [10], variability [22], and visualization [4, 20]. A vital fact to remark is, that if organizations strive for unleashing the potential of *BD*, it needs to be transferred into actionable output with the purpose to drive decision making — that is what the second component *analytics* is referring to [23, 10, 24].

*Analytics* manifests in various ways (e.g., data mining, text mining, machine learning, or PM) but always uses *BD* as the basis of its conduct [22]. This implies that the dimensions of *BD* represent a collection of characteristics for the data itself, *analytics* represents the tools, techniques, and analytical procedures to generate business value [10]. Thereby, the term BDA is focused on describing the output side [25] but it omits the explanation which inputs, in the form of resources, are required to shape an organizational asset [10].

To understand the purpose of BDAC and its

connection to organizational resources, it is essential to consider the theoretical background of its origin. The resource based view (RBV) [9] proposes an approach to disclose why the determined alignment and development of internal resources is the deciding factor for above-average firm performance and competitive advantage. The core essence of this theory is the development of resources towards a state where they become valuable (enable an organization to implement a value-creating strategy), rare (short in supply), imperfectly imitable (cannot be replicated by competitors), and non-substitutable (cannot be substituted by rivals)[9, 26].

The RBV was later specified by [27] to the extent that a distinction must be made between resource-picking and capability-building. While resource-picking is about identifying and controlling resources, capability-building is focused on orchestration and management of those resources with the purpose to evolve a valuable strategic asset [28, 27]. The difference between these two perspectives is, according to Amit and Schoemaker [29], the acquisition opportunity. While resources are tradable and not firm-specific, capabilities represent the opposite properties – a non-tradable and unique construct in the form of a high-level routine, that needs to be developed from within the organization [29, 10, 30, 31]. More specific, a capability is the result of the strategically aligned combination of many resources with the assistance of the own internal competence and knowledge [32, 33].

While resources can be easily replicated, a “distinctively set of capabilities mobilized by a firm” [3] is unique in its individual setting and can result in sustained competitive advantage [34, 35]. In order to shape a BDAC, the deciding factor is the right compilation and orchestration of many different detached resources that are available in the organization. Those identified core resources form the basis of the desirable BDAC. In this paper, three main domains of BDA resources which have been widely used in both literature and business context are utilized to describe those necessary resources – tangible resources (data, technology/infrastructure, and financing), intangible resources (data-driven culture and data governance), and human resources (data-specific technical skills and managerial competencies) [34, 10, 32, 13].

### 2.2. Process-oriented organizations

After generalized information about BDAC, this paragraph shifts the focus to a more process-oriented perspective to create the transition to the later elaborated

PM. After organizations created a BDAC they must ask themselves the upcoming question on which focus area they want to align their BDAC towards. For example, studies revealed that organizations from the media and news industry focus their BDAC activities towards “delivering tailored-made news and suggestions” [10], whereas companies from the oil and gas industry tend to concentrate more likely on “several applications geared towards risk assessment and maintenance” [10]. Nevertheless, every form of BDAC alignment results from a strategic decision and offers a certain type of value-adding component to the organization. In the case of process-orientation, the value-adding component is reflected in the form of improved performance, process efficiency, or process effectiveness [36, 2, 37].

Process-orientation indicates that an organization focuses its BDAC efforts mostly on business processes and their discovery, design, deployment, and execution [38]. By discovering, analyzing, and optimizing processes like that, organizations try to improve transparency and remove non-value adding activities more easily and therefore enhance profitability, customer satisfaction, time-to-market, and product or service quality [3, 38]. Nevertheless, Mikalef et al. [10] are mentioning that a process-oriented BDAC “is most likely to have significantly less impact on firm’s competitive position compared to utilizing them to detect new customer segments or come up with new business models”, although results can be measured more transparently and benefits are achieved in a faster and more tangible way [5, 39]. This gets even more important nowadays since “organizations no longer compete on processes [itself] but the ability to continually improve [them]” [38] to adapt to a fast-changing environment [40].

In order to now provide the opportunity to establish a connection from concrete organizational resources to the technical application of PM, the next section explains how to decompose PM in clearly designated configuration dimensions and maps them to their resource utilization.

### 2.3. Process mining configuration dimension setting

In the context of this paper, a *PM configuration dimension setting* (hereafter called setting) can be interpreted as a structure consisting of a set of five *PM configuration dimensions* classified in different categories. Based on [41] we defined those five categories as *field of application*, *perspective*, *usage type*, *online vs. offline*, and *context*. Adding to that, the limitation needs to be considered that each setting

never holds two configuration dimensions out of the same category since they are mutually exclusive. Each configuration dimension in turn implies a different facet of technical execution. Subsequently, each individual setting determines how to set up and execute PM from a technical, respectively algorithmic, point of view [41].

The category *field of application* holds the configuration dimensions discovery, conformance, and enhancement. Based on the underlying use case, it can be specified to uncover entirely new process models from an event-log (discovery), check discrepancies or irregularities of procedural models, organizational models or declarative process models (conformance) [42, 41]. Additionally, it is possible to improve existing process models by adding more detailed information to unveil “bottlenecks, service levels, throughput times, and frequencies” [41] (enhancement).

As the second category, the *perspective* takes into account which process-specific aspect of the data is analysed. In this respect, there is a classification between resource-oriented and performance-oriented as well as between specific and holistic [41]. As a holistic and performance-oriented perspective, the *control flow* focuses exclusively on the process activity layout. The *case* perspective can be interpreted as a specific and resource-oriented supplement to the control-flow perspective. It adds concrete properties to a case (e.g., the number of ordered products, the costs, or the supplier name) [41]. By using timestamp information from an event-log, the *time* perspective helps to understand execution time, latency, and the overall interrelation of activities to the chronological context [43]. Lastly, the holistic and resource-oriented *organizational* perspective examines the involvement of actors (e.g., people, systems, roles, or business units) [41] in conjunction to the actual process which can result in “social network graphs, assignment rules, and allocation constraints regarding which resources may execute” [43].

The *usage type* represents the third category with the intention to define the frequency and freedom of the PM usage. All three configuration dimensions of this category *standard*, *repeated*, and *ad-hoc* are classified as illustrated in Figure 1 [41].

Another distinction can be made by looking at the degree of case completion in the category *online vs. offline*. *Offline* only considers terminated process cases, respectively historical event data. On the contrary, *online* implies that process cases are analyzed during their execution to perform different kinds of analysis for detection, prediction, or recommendation purposes in real-time scenarios [41].

Finally, the category *context* defines the extent to

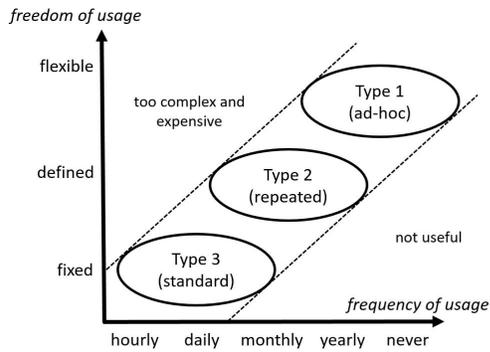


Figure 1. Usage types of PM [41]

which the process analysis is enriched with additional data from external origin. To highlight the importance, consider the fact that processes are subject to certain fluctuations due to daily, weekly, or seasonal patterns like “office hours, lunch breaks[, public holidays,] and weekends” [41] as well as factors attributable to the economic, political, cultural or ecological environment. Moreover, even different end-to-end processes inside the same organization, like order-to-cash in conjunction with procure-to-pay, influence each other because they mutually compete for the same organizational resources [44, 41]. With ascending complexity, there are the contexts *case*, *process*, *social*, and *external* [41].

As one might suppose, companies are often overwhelmed by the complexity of choosing the right PM configuration dimension settings in regards to their individual business problem — but even more difficult is the subsequent selection and acquisition of underlying BDA resources. Currently, companies are searching for decision support to cope with this challenge.

### 3. Research design

The research design described hereafter tries to antagonize the outlined research gap by assisting organizations with managing their PM-related activities toward building up their process-oriented BDAC.

As an overall framework for the research design, we use the design science research approach originated by Hevner et al. [15] which manifests itself in the form of an IT artifact, in this case, a PM alignment methodology, that repeatedly traverses two cycles in parallel several times until its completion — the behavioral science and design science cycle. Whereas the behavioral science cycle seeks for “development and justification of theories that explain or predict phenomena related to the identified business need[, the] [d]esign science [cycle] addresses research through the building and

evaluation of artifacts designed to meet the identified business need” [15]. The purpose of design science is associated with pragmatism and follows the goal to bridge science and practical action [45]. This is necessary to successfully cope with the difficulty to later implement the artifact in an organization and solve the tackled problems in a sustainable way, so that both, people who interact with the artifact as well as managers who want to improve effectiveness and efficiency are satisfied with the outcomes [46, 47].

The utilized design science research process includes the six activities “problem identification and motivation, definition of the objectives for the solution, design and development, demonstration, evaluation, and communication” [46]. The general design science research process by Peffers et al. [46], together with the actions we performed, is listed in Table 1. It represents the blueprint for the practical application of design science in this paper.

General step of design science process	Description of action
(1) Problem identification and motivation	Conduct and analyse semi-structured expert interviews with three-step procedure from [16].
(2) Define objectives of a solution	Derive the overall research goals directly from the identified first order problem space from step 1.
(3) Design and development	Develop a PM alignment methodology based on the goals from step 2.
(4) Demonstration	Pre-validate the developed methodology with so far uninvolved persons in regards to obvious errors and misleading usability by applying the methodology.
(5) Evaluation	Execute a comprehensive evaluation based on morphologically [48] justified sample vignettes [49].
(6) Communication	Publish the results to both practitioners and researchers.

Table 1. Steps from the design science research process [46]

Nonetheless, a vital part to remark is the identification of the problem space during its first activity. The difficulty lies in the clear demarcation and analysis of the problem space as well as the relevance and novelty of the problem that is unresolved [15, 46]. In conjunction with this particular research approach, the problem space has been identified and sharpened through semi-structured expert interviews which in turn have been structured and analyzed via a holistic approach of Gioia et al. [50]. This approach has

been chosen to improve the scientific significance and precision of inductive research by aggregating the statements and propositions of the interviewed sample data via a three-step procedure [50].

At the first level, all statements of the interviewed participants are collected unfiltered in their raw state and tokenized to get stand-alone statements of the first order. Duplicate statements with very similar meanings are combined to prevent the data basis from becoming too extensive. Hereinafter, the second level further increases aggregation by sorting the tokenized statements from the first level into general second-order themes. Special care is taken to ensure that this aggregation is based on the concepts presented in Section 2 and therefore target-oriented to the desired goal of sharpening the problem space. The final aggregation step breaks down the general themes of the second level into precise problem spaces (*aggregated dimensions*) which will serve as a starting point of the definition of objectives (second phase in Table 2) [16, 50].

## 4. Results

### 4.1. Problem identification and objectives

Before initiating the methodology creation process, it is inevitable to sharpen the problem space derived from the identified research gap that this methodology tries to address. Hevner et al. [15] mention this as one of the seven guidelines of design science by saying that an IT artifact which solves a non-existent business problem does not fulfill the requirements for being relevant in design science research. In order to identify and refine the specific problem space, this paper uses a qualitative approach by [50] described in Section 3. The metadata behind the conducted semi-structured expert interviews with consultants from the data & analytics competency of a Big-4 consultancy is listed in Table 2.

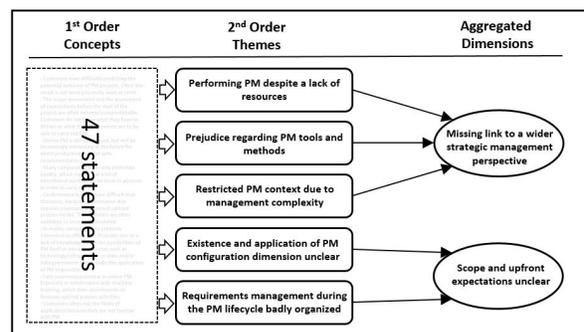
Rank	Focus sector	Professional experience	Interview length
Senior Manager	Life Science, Pharma	12 years	75 mins
Senior Manager	Automotive	10 years	37 mins
Senior Consultant	Automotive	6 years	55 mins
Senior Consultant	Life Science, Automotive	5 years	35 mins

**Table 2. Metadata from expert interviews to identify and analyse the problem space**

The interviews themselves follow a developed strand of action which is based on the literature presented earlier. In support of that, all interviewed participants

were given a pre-prepared guideline in advance on the basis of which the subsequent interview was then executed. Overall, there were five general questions about potential problems, challenges, and potentials of the various configuration dimensions of PM. Important to add, respondents were asked to refer directly as their answers to the organizational BDA resources. To clarify that, the purpose of this structure helps to later connect the collected statements of the problem space to the knowledge base with the goal to diminish the potential rigor-relevance gap [51].

The outcome of the interview phase generated 47 first-order concepts which were then conflated into five second-order themes. The two second-order themes *existence and application of PM configuration dimensions unclear* and *requirements management during the PM life cycle badly organized* resulted in the first aggregated dimension *scope and upfront expectations unclear*. Whereas the remaining three second-order themes *performing PM despite a lack of resources*, *prejudice regarding PM tools and methods*, and *restricted PM context due to management complexity* coalesced into the second aggregated dimension *missing link to a wider strategic management perspective* (Figure 2).



**Figure 2. Identified problem spaces**

This subsection closes with the fixation of qualitative objectives as the second activity from the design science research process (Table 1). The goals directly address the previously identified problem spaces from Figure 2 and underpin them with concrete actions for the design phase:

- Address first problem space *scope and upfront expectations unclear* by means of a target-oriented method of requirements analysis for PM projects which offers a fast and easy to handle usability. The approach to be developed should be flexible and adaptable to almost all existing PM constellations. Moreover, the approach should be independent of particular

vendors or tools in order to give the best possible solution.

- Address second problem space *missing link to a wider strategic management perspective* by linking the discovered PM configuration dimension settings to the organizational resource perspective. Additionally, it should be possible to identify which organizational resources are needed and how to orchestrate them to gain a process-oriented BDAC.

#### 4.2. Design and development

To cover up the previously defined objectives, a structured and comprehensive methodology is created – it represents the PM alignment methodology which internalizes the result of this particular design science approach. Overall, there have been two iteration cycles which final result is elaborated hereafter. However, before that, it is essential to impart the theoretical reasoning behind the methodology. The intention of the developed methodology is to bring PM, as a BDA technique, together with the concept of BDAC by indicating how it can support the alignment towards process orientation. Therefore, the PM component is directed towards two key positions as illustrated in Figure 3.

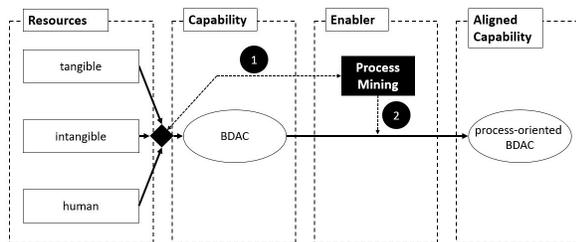


Figure 3. Presumption of interrelations of the designed PM alignment methodology (adapted from [13, 10])

Firstly, it supports during the ex-ante selection and orchestration of BDA resources needed for effective PM execution (Number 1 in Figure 3). This is achieved by a determined requirements analysis about the degree of how important specific PM configuration dimensions are for the organization. With this knowledge in mind, it is possible to select the right organizational resources as well as in which proportion they need to be weighted (e.g., one specific PM configuration dimension setting demands real-time availability with highly complex event-log data). In this case, especially the BDA resources *infrastructure/technology* and *data-specific*

*technical skills* are essential and need to be developed or acquired inside the organization.

Following the ex-ante resource selection and orchestration, the second key position of PM is getting relevant (Number 2 in Figure 3) because so far, the BDAC is only established but not aligned. What is missing to achieve a value-generating alignment is the actual application of PM in conjunction with the selected BDA resources since no process is optimized until this point of time [5, 32]. This is where PM comes in again as an enabler because now organizations can use the previously specified setting ex-post with its related BDA resources to perform PM on their processes in order to enable continuous optimization. Due to the existence of the BDAC, they can accomplish this in a more effective and efficient manner.

The developed methodology is implemented on the basis of Microsoft Excel and is intended to be a sequential approach, which is closely based on the theoretical foundations. Overall, it consists of five tabs related to the purpose to capture PM-specific requirements as well as a summary tab which is intended to give the final recommendation of what configuration dimension setting is most suitable for a certain scenario but also which BDA resources are needed for the subsequent execution. For administrative use only, two additional tabs are added to the Excel tool in order to make the individual configurations for the resource assessment, calculation weightings, KPI selection as well as vendor characteristics.

Category	Specification	Preferred scenario	Relative Score	Absolute Score
Field of application	Discovery	I want	46,15384615	46,15384615
		My int	45,71428571	45,71428571
	Conformance	My pri	35,8974359	35,8974359
		The pi	36,666667	36,666667
	Enhancement	I want	17,94871795	17,94871795
		I am a	15,33333333	15,33333333

Figure 4. Screenshot of the tab 'Field of application' of the PM alignment methodology

The first five tabs are assigned directly to the previously mentioned categories from Section 2.2, namely *field of application*, *perspective*, *usage type*, *online vs. offline*, and *context*, and collectively follow the structure of Figure 4. The purpose of all five tabs is to capture the requirements of the needed

PM configuration dimension setting and the coherent selection and orchestration of the BDA resources. Section one in Figure 4 contains the scoping statements area. For each PM configuration dimension of a category ( $i \in M$ ), three statements are provided to narrow down the needs of the user questioned. To be more specific, the user can determine the level of approval regarding the three statements above via a seven-unit Likert scale ( $x_{i1}, x_{i2}, x_{i3}$ ). On the left side of the scoping statement area is the scoring section. This scoring functionality automatically calculates the most suitable configuration dimension of a category based on the selected value on the Likert scale and transfers this into an ordinal order. For instance, Figure 4 suggests that *discovery* has the highest priority within the category *field of application* with a score of 80.95% in absolute ( $a_i$ ) and 41.46% in relative ( $r_i$ ) numbers. Following the evaluation of the statements, the user can press the *Go to next step* button (Number 3 in Figure 4) to move to the next category until all five categories are completed.

$$a_i = \frac{100}{7} * \left( \frac{x_{i1} + x_{i2} + x_{i3}}{3} \right) \quad (1)$$

$$r_i = \frac{100}{\sum_{i \in M} a_i} * a_i \quad (2)$$

As a result of the five previous steps, the *summary* tab collects all information and displays it in an overview (Figure 5). Recognizably itemized in the second row of this screenshot is the final PM configuration dimension setting – in this case, an example with *discovery*, *time*, *standard*, *offline*, and *external*. This selected setting is based on the maximum relative value from each category tab ( $r_i$ ). Furthermore, this tab offers the first point of contact to an organizational perspective since directly underneath the final configuration dimension setting, the associated intensity of the required BDA resources are outlined. To be more precise, concrete values state to what extent a specific BDA resource is required to perform the related setting (which can be predefined and weighted in the *setting* tab).

The illustrated horizontal bar diagram in the middle of Figure 5 indicates the relation of required BDA resources for the whole setting. The percentage values indicate in which relation a certain BDA resource is needed in comparison to the other BDA resources (e.g., a high value of the intangible BDA resource *data-driven culture* results from the suggested PM configuration dimensions *online* and *external*). By considering this, the user gets an orientation of how to select and orchestrate the resources in order to achieve the best-possible PM execution. To be clear, the resulting

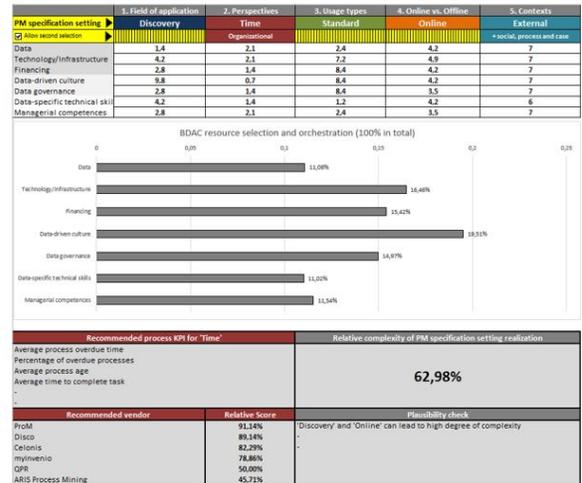


Figure 5. Screenshot of the tab 'Summary' of the PM alignment methodology

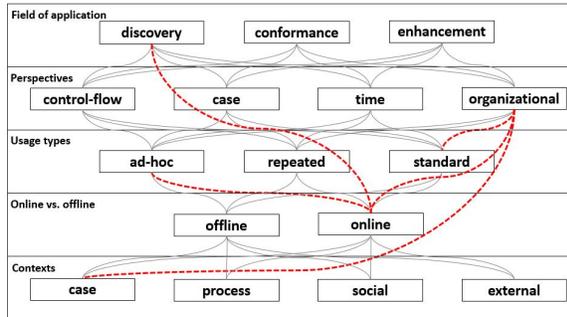
percentage weighting is characterized by considerable uncertainty since a lot of contextual information of the organization cannot be captured. When reminiscing Figure 3, it becomes evident that this is related to the first touch point from the theoretical concept behind this methodology.

In addition, other implications can be derived from the *summary* like an indication of the overall complexity of the PM configuration dimension setting, KPI suggestions, a vendor suggestion as well as a plausibility check which is based on the morphological analysis during the evaluation phase.

### 4.3. Evaluation and demonstration

In order to evaluate the designed methodology, we use three different vignettes to get the desired information about how accurate the problem spaces are being addressed [52, 49]. All developed vignettes are based on conducted PM consulting projects and therefore fit with typical scenarios associated with large multinational companies that are aiming at the optimization of entire (standard) end-to-end processes. The vignettes assume that PM is the preferred technique for problem-solving. The basis for the vignette development is a morphological analysis of possible constellations of PM configuration dimension settings since although every constellation is possible, not all of them make sense for the practical application. Figure 6 illustrates this procedure by using red dotted lines for unfavorable settings (e.g., the *organizational* perspective is not likely to be used as a *standard* usage type or in the very limited *case* context) and grey solid

lines for purposeful settings (e.g., it is well conceivable to use the *control-flow* perspective in conjunction with an *ad-hoc* usage type). A particular vignette manifests itself as a short concise prose text that implicitly contains one dedicated setting out of all previously uncovered possible constellations.



**Figure 6. Morphological analysis of likely/unlikely vignette constellations**

Overall, two design science iterations were conducted. Equivalent to that, every design and development activity was followed by an evaluation. Both iterations referred to the same three vignettes constellations to get comparable results and examine whether an improvement had occurred. All vignettes were then sent to the same four experts who already participated in sharpening the problem spaces together with a related questionnaire, the PM alignment methodology itself, and a short explanation about its practical application. In this regard, the respondents had to put themselves into a role and apply the methodology practically on the basis of the vignettes. Usability, functionality, completeness, consistency, and fit with the organization were some of the evaluation criteria suggested by [15] which were also adopted by this questionnaire. The assumed benchmark was to reach an overall consistency alignment between the intended PM configuration dimension setting behind the vignette and the actual results from the questionnaire of at least 80% (1. iteration: 74.94% / 2. iteration 81.64%). By achieving the goal, the design science process was able to proceed to its final step *communication*.

## 5. Discussion, communication, and limitations

When contemplating about the term *BD* or *analytics* in general, processes are usually not instantaneously associated since the focus is more on new product development or exploring customer behavior [10, 2]. This neglect is also reflected in the literature, as there

is relatively little research on strategically considering BDA (in particular its manifestation in the form of PM) as a crucial enabler for competitive advantage by leveraging data-driven process optimization approaches. The results from the empirically conducted problem space identification confirm those assumptions. The majority of current research contributions measures the impact of BDAC on firm performance [2, 3, 32] or on competitive advantages but none investigates how to evolve a (process-oriented) BDAC and sustain or adapt it over a certain period of time [4, 13, 53]. Also, a vast majority of the BDAC contributions concentrate on the resource-picking aspect but neglect the linkage to clear advice on which activities to follow and techniques to use [10]. On the contrary, research in the field of PM is dominated by its technical usage in terms of hands-on topics like algorithms, operational project management, or associated infrastructure [41, 43] but it disregards referencing to organizational theories.

The design science approach presented in this paper delivers two kinds of IS research contributions. On the one hand, it discloses the connection between PM as an analytics technique and the organizational concept of process-oriented BDAC. It represents the first research attempt of this kind to actively use PM as an enabler to achieve strategical objectives with organizational alignment. Thereby, it enriches the research in the area of BPM capability alignment [54, 55] by focusing more on its data-driven perspective. On the other hand, it consequently transmits this theoretical construct into a PM alignment methodology, which can be used in real-world scenarios for executing this organizational alignment towards process-orientation. To be more precise, it advises on the PM requirements phase by helping to capture the whole spectrum of possibilities just to transform these findings into the demanded BDA resources. In addition to the purely content-related contributions, this paper supplements the design science community by rigorously applying its seven guidelines [15] on a different field of application than its popularly used ones (e.g., maturity models or workflow/process language development) [56, 57, 58].

Like any other research, this study is not without limitations. In case of the problem identification, the breaking down of second-order themes to an aggregated problem dimension creates a loss of information and sets limits to generalizability. Nevertheless, this was necessary to establish a clear link between the statements of the interviewees and overarching problem spaces. The vignettes applied during the evaluation are of a broadly formulated nature without a focus on company size, industry sector, or firm structure, although this could have impacted the findings notably.

Additionally, the settings are fictitious and relatively sensitive to textualization.

## 6. Conclusion

In this paper, we address the question of how BDAC can be strategically aligned with PM. To this end, we designed a PM alignment method that helps companies mapping their individual technical requirements of PM to associated organizational resources. We evaluated our method using a series of interviews with IT consultants. Our research is a first contribution in the domain of BDA that discusses PM beyond technical and operational questions of project management. In this way, it provides a foundation for future research on the business implications of PM.

## References

- [1] M. Dumas, M. L. Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management, Second Edition*. Springer, 2018.
- [2] S. Wamba, S. Akter, A. Edwards, G. Chopin, and D. Gnanzou, "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study," *International Journal of Production Economics*, vol. 165, pp. 234–246, 2015.
- [3] S. F. Wamba, A. Gunasekaran, S. Akter, S. J.-f. Ren, R. Dubey, and S. J. Childe, "Big data analytics and firm performance: Effects of dynamic capabilities," *Journal of Business Research*, vol. 70, pp. 356–365, 2017.
- [4] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of big data challenges and analytical methods," *Journal of Business Research*, vol. 70, pp. 263–286, 2017.
- [5] G. Kim, B. Shin, K.-K. Kim, and H.-G. Lee, "It capabilities, process-oriented dynamic capabilities, and firm financial performance," *Journal of the Association for Information Systems*, vol. 12, no. 7, pp. 487–517, 2011.
- [6] J. vom Brocke and J. Mendling, eds., *Business Process Management Cases, Digital Innovation and Business Transformation in Practice*. Management for Professionals, Springer, 2018.
- [7] S. Aguirre, C. Parra, and M. Sepúlveda, "Methodological proposal for process mining projects," *International Journal of Business Process Integration and Management*, vol. 8, no. 2, pp. 102–113, 2017.
- [8] M. L. van Eck, X. Lu, S. J. Leemans, and W. M. van der Aalst, "Pm2: A process mining project methodology," in *International Conference on Advanced Information Systems Engineering*, pp. 297–313, Springer, 2015.
- [9] J. Barney, "Firm resources and sustained competitive advantage," *Journal of Management*, vol. 17, no. 1, pp. 99–120, 1991.
- [10] P. Mikalef, V. A. Framnes, F. Danielsen, J. Krogstie, and D. Olsen, eds., *Big Data Analytics Capability: Antecedents and Business Value*, vol. 136, PACIS 2017, 2017.
- [11] M. Peteraf, G. Di Stefano, and G. Verona, "The elephant in the room of dynamic capabilities: Bringing two diverging conversations together," *Strategic Management Journal*, vol. 34, no. 12, pp. 1389–1410, 2013.
- [12] S. Akter, S. F. Wamba, A. Gunasekaran, R. Dubey, and S. J. Childe, "How to improve firm performance using big data analytics capability and business strategy alignment?," *International Journal of Production Economics*, vol. 182, pp. 113–131, 2016.
- [13] P. Mikalef, I. O. Pappas, J. Krogstie, and M. Giannakos, "Big data analytics capabilities: a systematic literature review and research agenda," *Information Systems and e-Business Management*, vol. 17, no. 2, p. 1, 2017.
- [14] R. Cosic, G. Shanks, and S. Maynard, eds., *Towards a Business Analytics Capability Maturity Model*, (Geelong), 23rd Australasian Conference on Information Systems, 2012.
- [15] A. Hevner, A. R. S. March, S. T. Park, J. Park, Ram, and Sudha, "Design science in information systems research," *Management Information Systems Quarterly*, vol. 28, no. 1, pp. 75–105, 2004.
- [16] D. A. Gioia, "A renaissance self: Prompting personal and professional revitalization," in *P. J. Frost & R. E. Stablein (Eds.), Renewing research practice: Scholars' Journeys*, pp. 97–114, 2004.
- [17] M. Rosemann and J. vom Brocke, "The six core elements of business process management," in *Handbook on business process management 1*, pp. 105–122, Springer, 2015.
- [18] M. Berner, E. Graupner, and A. Maedche, "The information panopticon in the big data era," *Journal of Organization Design*, vol. 3, no. 1, p. 14, 2014.
- [19] A. McAfee and E. Brynjolfsson, "Big data: The management revolution," *Harvard business review*, vol. 90, 2012.
- [20] P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," *Information Sciences*, vol. 275, pp. 314–347, 2014.
- [21] M. White, "Digital workplaces," *Business Information Review*, vol. 29, no. 4, pp. 205–214, 2012.
- [22] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International Journal of Information Management*, vol. 35, no. 2, pp. 137–144, 2015.
- [23] M. M. Gobble, "Big data: The next big thing in innovation," *Research-Technology Management*, vol. 56, no. 1, pp. 64–67, 2013.
- [24] Z. Xu, G. L. Frankwick, and E. Ramirez, "Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective," *Journal of Business Research*, vol. 69, no. 5, pp. 1562–1566, 2016.
- [25] H. S. Lamba and S. K. Dubey, *Analysis of requirements for big data adoption to maximize IT business value in reliability*. In: 2015 4th International conference on infocom technologies and optimization (ICRITO): IEEE, pp 1-6, 2015.
- [26] A. Schwarz, M. Kalika, H. Kefi, and C. Schwarz, "A dynamic capabilities approach to understanding the impact of it-enabled businesses processes and it-business alignment on the strategic and operational performance of the firm," *Communications of the Association for Information Systems*, vol. 26, no. 4, 2010.

- [27] R. Makadok, "Toward a synthesis of the resource-based and dynamic-capability views of rent creation," *Strategic Management Journal*, vol. 22, no. 5, pp. 387–401, 2001.
- [28] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity," *McKinsey Global Institute*, vol. Seattle, 2011.
- [29] R. Amit and P. J. H. Schoemaker, "Strategic assets and organizational rent," *Strategic Management Journal*, vol. 14, no. 1, pp. 33–46, 1993.
- [30] D. J. Teece, G. Pisano, and A. Shuen, "Dynamic capabilities and strategic management," *Strategic Management Journal*, vol. 18, no. 7, pp. 509–533, 1997.
- [31] K. M. Eisenhardt and J. A. Martin, "Dynamic capabilities: what are they?," *Strategic Management Journal*, vol. 21, no. 10–11, pp. 1105–1121, 2000.
- [32] M. Gupta and J. F. George, "Toward the development of a big data analytics capability," *Information & Management*, vol. 53, no. 8, pp. 1049–1064, 2016.
- [33] S. A. Zahra and A. P. Nielsen, "Sources of capabilities, integration and technology commercialization," *Strategic Management Journal*, vol. 23, no. 5, pp. 377–398, 2002.
- [34] T. H. Davenport, P. Barth, and R. Bean, "How 'big data' is different," *Sloan Management Review*, vol. 54, 2012.
- [35] M. Wade and J. Hulland, "Review: The resource-based view and information systems research: Review, extension, and suggestions for future research," *MIS Quarterly: Management Information Systems*, vol. 28, no. 1, pp. 107–142, 2004.
- [36] R. Bärenfänger, B. Otto, and H. Österle, "Business value of in-memory technology – multiple-case study insights," *Industrial Management & Data Systems*, vol. 114, no. 9, pp. 1396–1414, 2014.
- [37] R. Škrinjar, V. Bosilj Vukšić, and M. Indihar-Štemberger, "The impact of business process orientation on financial and non-financial performance," *Business Process Management Journal*, vol. 14, no. 5, pp. 738–754, 2008.
- [38] M. Kohlbacher, "The impact of dynamic capabilities through continuous improvement on innovation: the role of business process orientation," *Knowledge and Process Management*, vol. 20, no. 2, pp. 71–76, 2013.
- [39] F. Damanpour and S. Gopalakrishnan, "The dynamics of the adoption of product and process innovation in organizations," *Journal of Management Studies*, vol. 38, no. 0022-2380, pp. 45–65, 2001.
- [40] D. J. Teece, "Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance," *Strategic Management Journal*, vol. 28, no. 13, pp. 1319–1350, 2007.
- [41] W. van der Aalst, *Process Mining - Data Science in Action*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2 ed., 2016.
- [42] W. M. P. van der Aalst, V. Rubin, H. M. W. Verbeek, B. F. van Dongen, E. Kindler, and C. W. Günther, "Process mining: a two-step approach to balance between underfitting and overfitting," *Software & Systems Modeling*, vol. 9, no. 1, pp. 87–111, 2010.
- [43] F. Mannhardt, *Multi-perspective process mining*. PhD thesis 1 (research tu/e / graduation tu/e), Technische Universiteit Eindhoven, Eindhoven ER -, 2018.
- [44] D. Schahram and W. van der Aalst, "Process mining put into context," *IEEE Computer Society*, vol. 12, pp. 82–86, 2012.
- [45] W. J. Orlikowski and J. J. Baroudi, "Studying information technology in organizations: Research approaches and assumptions," *Information Systems Research*, vol. 2, no. 1, pp. 1–28, 1991.
- [46] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007.
- [47] J. Iivari, "Twelve theses on design science research in information systems," in *Design Research in Information Systems* (A. Hevner and S. Chatterjee, eds.), vol. 22 of *Integrated Series in Information Systems*, pp. 43–62, Boston, MA: Springer US, 2010.
- [48] V. Kayser, *Web-based Scenario Development - Process Improvements*. PhD thesis, TU Berlin, Berlin, 2016.
- [49] J. Cummings and A. Dennis, "Virtual first impressions matter: The effect of enterprise social networking sites on impression formation in virtual teams," *MIS Quarterly: Management Information Systems*, vol. 42, pp. 697–718, 09 2018.
- [50] D. A. Gioia, K. G. Corley, and A. L. Hamilton, "Seeking qualitative rigor in inductive research," *Organizational Research Methods*, vol. 16, no. 1, pp. 15–31, 2013.
- [51] K. Starkey and P. Madan, "Bridging the relevance gap: Aligning stakeholders in the future of management research," *British Journal of Management*, vol. 12, no. s1, pp. S3–S26, 2001.
- [52] A. S. Silva, W. L. Campos-Silva, M. A. Gouvea, and M. C. Farina, "Vignettes: a data collection technique to handle the differential operation of items in surveys," *BBR. Brazilian Business Review*, vol. 16, pp. 16 – 31, 02 2019.
- [53] M. J. Mazzei and D. Noble, "Big data dreams: A framework for corporate strategy," *Business Horizons*, vol. 60, no. 3, pp. 405–414, 2017.
- [54] B. Niehaves, J. Poepfelbusch, R. Plattfaut, and J. Becker, "Bpm capability development – a matter of contingencies," *Business process management journal*, vol. 20, 09 2014.
- [55] M. Rosemann and J. vom Brocke, *The Six Core Elements of Business Process Management*, pp. 105–122. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015.
- [56] M. Röglinger, J. Pöppelbuß, and J. Becker, "Maturity models in business process management," *Business Process Management Journal*, vol. 18, no. 2, pp. 328–346, 2012.
- [57] J. Lee, D. Lee, and S. Kang, "An overview of the business process maturity model (bpmm)," in *Advances in Web and Network Technologies, and Information Management* (K. C.-C. Chang, W. Wang, L. Chen, C. A. Ellis, C.-H. Hsu, A. C. Tsoi, and H. Wang, eds.), vol. 4537 of *Lecture Notes in Computer Science*, pp. 384–395, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
- [58] W. M. P. van der Aalst and A. Kumar, "Xml-based schema definition for support of interorganizational workflow," *Information Systems Research*, vol. 14, no. 1, pp. 23–46, 2003.