

Designing a Process Mining-Enabled Decision Support System for Business Process Standardization in ERP Implementation Projects

Christian Fleig¹, Dominik Augenstein¹, and
Alexander Maedche¹

¹Karlsruhe Institute of Technology (KIT),
Institute of Information Systems and Marketing (ISM), Karlsruhe, Germany
christian.fleig@kit.edu

Abstract. Process standardization allows to optimize ERP systems and is a necessary step prior to ERP implementation projects. Traditional approaches to standardizing business processes are based on manually created "de-jure" process models, which are distorted, error-prone, simplistic, and often deviating from process reality. Theoretically embedded in the organizational contingency theory as kernel theory, this paper employs a design science approach to design a process mining-enabled decision support system (DSS) which combines bottom-up process mining models with manually added top-down standardization information to recommend a suitable standard process specification from a repository. Extended process models of the as-is process are matched against a repository of best-practice standard process model using an attribute-based process similarity matching algorithm. Thus, the DSS aims to reduce the overall costs of process standardization, to optimize the degree of fit between the organization and the implemented processes, and to minimize the degree of organizational change required in standardization and ERP implementation projects. This paper implements a working prototype instantiation in the open-source process analytics platform Apromore based on a real-life event log and standardization attributes for the Purchase-to-Pay and Order-to-Cash processes from three SAP R/3 ERP systems at the industry partner.

Keywords: Process Mining, Decision Support Systems, Process Standardization, Enterprise Resource Planning, ERP Implementation Projects

1 Introduction

Rapidly evolving competitive environments and emerging business opportunities require the transformation of business processes in the organization in response to new conditions to remain competitive [1]. However, the transformation of a business process from a current design to a target process design requires organizations to precisely understand the real-world execution of the as-is process to make solid transformation decisions (e.g., [2]). Organizations frequently do not meet these prerequisites for business process transformation, and possess only limited insights and a narrow understanding of existing process execution paths [3]. Traditional top-down approaches to

business process transformation rely on "de-jure" process analyses instead of bottom-up "de-facto" data-driven approaches. These "de-jure" approaches suffer from a number of insufficiencies as they are based on handmade process models which are often biased compared to process reality [4]. "De-jure" process documentations usually only contain idealistic process executions such as the to-be process, while most process variants and deviations from the ideal target specification are ignored [5]. In addition to content-related insufficiencies, top-down process modeling itself is a time- and resource-consuming task [6]. Further, "de jure" process models are error-prone due to their manual creation. In sum, van der Aalst finds that the currently prevailing approaches of process modelling are "disconnected" from process realities [7], which implies that human-centered top-down approaches provide only an insufficient base for decision-making in process transformation.

A chance to overcome these weaknesses of decision-making in process transformation is to utilize the increasing availability of process data from numerous information sources in organizations [8]. For example, information systems store process events in large event log tables [9] which provides the possibility to improve decision-making by data-driven approaches to process analytics such as process mining [5]. For example, process mining delivers descriptive and positive "de-facto" process analyses based on bottom-up data [5]. Hence, "de-facto" process analyses provide a valuable complement to decision-making in process transformation.

As a particular field of process transformation, business process standardization has experienced a high degree of scholastic attention [10], and has been recognized as a critical step prior to the implementation of new enterprise resource planning (ERP) systems (e.g., [11]). However, ERP systems such as SAP or Oracle provide numerous alternatives of possible standard processes by software vendors. In workshops performed at our industry partner in the context of a large-scale business process standardization and SAP S/4 HANA implementation project, we discovered that organizations are frequently challenged by the selection of the most appropriate standard process design. Thus, organizations might significantly benefit from a decision support systems (DSS) in the selection of suitable standard processes which overcomes the outlined weaknesses, and which considers the very specific process requirements of the individual organization. The research question of this paper therefore becomes:

How to design a process mining-enabled decision support system to support organizations in the standardization of business processes?

Besides a practical need for data-driven standardization decisions, an important research gap refers to the absence of contributions on the "post-mining" phase, with only few contributions exploring the question of how to turn the insights gained by process mining into actual process transformation decisions. This paper employs a design science research (DSR) approach to motivate, conceptualize, develop, and to evaluate the DSS artifact.

The rest of this paper is organized as follows. Section 2 briefly introduces conceptual foundations. Section 3 describes the design science research (DSR) methodology to systematically derive and implement a working DSS in the Apomore process analytics platform [12]. Section 4 derives meta-requirements (MRs) for the DSS which serve as developmental guidelines. Section 5 further concretizes the MRs in design principles (DPs) and design decisions (DDs) to develop the technical blueprint conceptualization. Section 6 describes the implementation in Apomore based on data from three real-

world SAP R/3 ERP systems for the purchase-to-pay (“Purchasing”) and order-to-cash (“Sales”) processes from a manufacturing corporation. Section 7 presents results from three evaluations of different aspects of the DSS. Section 8 concludes and presents limitations and avenues for future research.

2 Related Work

This section lays the literature foundations for the design requirements to the process mining-enabled DSS. To achieve the intended purpose, the DSS requires theoretical embedding in literature on process standardization, process mining, and similarity-based process matching. This section introduces the kernel theory from organization science as theoretical embedding and presents related work for the derivation of meta-requirements and design principles and decisions in section 4 and section 5.

This research is theoretically motivated by organizational contingency theory by Donaldson [13] and Sousa and Voss [14], which requires organizations to adapt structures to maintain a fit between changing contextual factors and environmental variables to retain performance. With particular regard to the field of BPM, as business processes are highly context-dependent (e.g., [15]) and business processes are systems which interact with the environment [16]. Extant research such as the contribution by vom Brocke et al. finds the effect of process management to be contingent upon contextual factors including organizational factors, process characteristics, and goals [17]. Thus, contingency theory requires a fit between business processes and environments [18], and to adapt business processes in response to any change in environmental variables. In addition to generic contingency theory, the organizational information processing theory by Galbraith [19] considers organizations as information-processing units which collect and process information and thus need to fit variables inside and outside the organization [20]. Therefore, the DSS designed in this research further incorporates contextual process factors such as standardization attributes to yield standardization support based on the contingencies of a particular organization.

We define a DSS as any system to address semi-structured or unstructured problems to support decision-making processes of users (e.g., [21]). Besides, Numerous contributions reveal a vital importance of alignment between the organization, business processes, and ERP systems [22]. Process standardization aims at a situation where the same activity in different organizational units is performed identically [23]. A standardized process “is constantly performed following the same steps in the same sequence” [24] and standardization can be achieved by the application of formalities, e.g. by creating guidelines or work instructions [24].

Contemporary information systems such as WfM, ERP, CRM, SCM, and B2B systems record business events in so-called event logs, which serve as foundations for process mining [3, 9]. For example, SAP logs all transactions, e.g., users filling out forms, changing documents etc. which significantly improves the ability to derive process transformation decisions by taking into account process variants and additional process information. To overcome the outlined weaknesses of top-down approaches to process standardization, this paper aims to utilize process to select standard processes using process mining as a source of bottom-up process information. Process mining aims to automatically discover business processes from transaction data [9, 25], and

offers a spectrum of techniques to perform automatic process discovery, monitoring, and improvement activities using system data in event logs [4]. In particular, process mining retrieves process models, which graphically and analytically represent business processes [22] and depict the course of activities and their dependencies [26].

In addition to process mining, the DSS is required to perform a matching of the as-is process against best-practice standard processes to propose a suitable standard process for implementation. The application of similarity for process matching is motivated by the minimization of disruptiveness of the new future process design to the organization. ERP implementation projects impose a “technochange” situation on organizations as ERP projects simultaneously impact technological as well as organizational structures. Technochange situations require significant efforts in terms of IT project management and change management [22]. Hence, adequate to-be standard processes are characterized by a high degree of implementability. Implementability addresses limitations in the organizational adaptability, and thus requires a minimum of misfits to the organization [27]. Therefore, selecting business process designs X' which exhibit a high degree of similarity to the current as-is process in (X) reduces misfits of the selected and the status-quo business process. Misfits are the result of low similarity between the current business process and the future business process. The resulting transformation for a business process with a low level of similarity between X and X' requires large transformation efforts, which overhauls routines and modifies well-acustomed workflows. As a consequence, adverse technochange situations and risks might arise for the organization such as high costs, a reduction of organizational performance, or the avoidance of the information system when choosing a target process with a low degree of similarity to the as-is process.

3 Research Methodology

Design science develops artifacts to address important organizational problems [28]. Thus, this paper employs a design science research (DSR) approach to provide organizations with a “process mining-enabled DSS” in two design cycles in a “build-and-evaluate-loop” [28]. In addition to providing a software artifact, we aim to derive the design requirements as a theoretical contribution for the system to abstract from the concrete artifact. We conduct the DSR project within the context of a large-scale ERP implementation project, which comprises the replacement of the current SAP R/3 ERP by the future SAP S/4 HANA Business Suite. In 2017, the corporation consisted of several sub-companies operating globally with more than 8.200 employees and about 1.2bn Euro in turnover. The industry partner provided an event log for the purchase-to-pay (“Purchasing”) and the order-to-cash (“Sales”) process for three companies for period from 01/2016 to 07/2017. Therefore, this contribution therefore uses a real-life event log, and thus overcomes the weaknesses of many process mining contributions when relying on synthetic, simulated data. Project responsibility is allocated to a coordination team of senior decision-makers which serve as workshop participants to derive meta-requirements and design principles apart from literature. Each design cycle consists of a problem awareness, a suggestion, a development, and an evaluation phase as proposed by the seminal contribution by Hevner et al. [28]. In the problem awareness phase of cycle one, a structured literature on process mining and ERP implementation

projects was conducted to validate the theoretical research gap and the need for decision support in business process standardization in the context of ERP implementation projects. An important gap in process mining research is characterized by the lack of research on the "post-mining phase", with almost no contributions investigating the question of how the findings from process mining can actually be used in the standardization of business processes to support ERP implementations. In the suggestion phase of the DSR project to address the research problem, meta-requirements and design principles are derived in four workshops. Participants include decision-makers from the different sub-companies, namely the chief information officer (CIO), the project leader (IT / ERP process expert), a leading operations manager (manager supply chain execution), a sales process expert (supervisor market research), a senior accountant (director controlling), an external IT and ERP consultant, and the PhD-student (first author of this paper in a passive form). Further, a literature review was conducted to enrich meta-requirements from practitioner workshops with theoretical foundations from the fields of process standardization, process mining, ERP implementations, and process matching techniques. In the development phase of cycle one, a DSS prototype was developed in Apromore [12]. As the entire system can hardly be evaluated in a single evaluation, the evaluation is split into different evaluations of individual system aspects. The first design cycle performs an evaluation of three aspects of the system. First, a technical evaluation is performed to demonstrate feasibility of the approach. Second, the system links process models with standardization attributes. These attribute-extended process models are evaluated in terms of the ability to increase process model comprehension of decision-makers. Third, semi-structured interviews are performed with decision-makers to determine system quality and usefulness in process standardization (e.g., [29]).

Design cycle two will consist of a refinement of meta-requirements and design principles to arrive at a final conceptualization of the DSS. The second design cycle will further concretize the design requirements to incorporate learnings from the evaluations performed in cycle one and to improve the Apromore artifact. In particular, following the demonstration of technical feasibility with real data in the Apromore instantiation, solid design science requires a further evaluation of the process model matching algorithm in future research. Findings in the evaluation from the previous design cycle will be implemented in the DSS instantiation to finalize the Apromore software artifact. Design cycle two will close with a second evaluation round in terms of whether managers would actually decide to adopt the DSS in projects.

4 Meta-Requirements

Organizational process knowledge might either be stored in prescriptive "to-be" and top-down sources of information such as the implicit knowledge of process participants or be stored in descriptive "as-is" bottom-up sources such as information systems. As each of these two types of process information has individual strengths and weaknesses, the DSS needs to be able to retrieve and combine process knowledge from different sources, and to combine these different types of process information before deriving decision support for business process standardization.

Thus, the DSS needs to incorporate bottom-up process information into decision-making to provide "as-is" process-specific standardization guidance. A potential source

is bottom-up process information stored in information systems such as ERP systems. These sources include data generated by systems during process execution, such as event log tables within the ERP systems.

MR1: The DSS needs to incorporate de-facto bottom-up process information.

These data sources capture process executions “as-is”. An exclusive reliance upon process mining in decision-making for business process standardization yields merely an incomplete picture of process realities. Process mining captures only information on process activities within the information system (e.g., [4]), and event logs merely contain a subset of all possible process facets [4, 5]. Therefore, insights gained from bottom-up sources might be incomplete due to shadow process steps which are not recorded in the system event log. The DSS needs to incorporate different types of quantitative and qualitative process information in addition to bottom-up models and additional process knowledge needs to be retrieved from top-down sources. In particular, top-down sources comprise intangible human process knowledge which cannot be retrieved bottom-up as these process elements are not executed within the information system. Examples include paper-based process steps, third-systems, inputs, outputs, off-system data, and participating user groups. We introduce MR2 accordingly:

MR2: The DSS needs to provide a user interface to retrieve additional top-down process information.

Organizational contingency theory by Donaldson [13] requires activities of business process management to consider the respective circumstances and contexts of business processes into decision-making. The work by Rosemann and Vessey [30] introduces the notion of context-dependent processes. Therefore, the DSS needs to incorporate relevant contextual process information and to capture information such as standardization goals, process type, and key process dimensions and characteristics to provide tailored decision support depending on the circumstances of the respective organization and the respective business process. We consequently introduce MR2a:

MR2a: The DSS needs to incorporate process context factors and process characteristics into decision-making.

Furthermore, most approaches in BPM usually incorporate strategic process goals [31] which are compatible with the overall organization strategy. These transformation goals serve as an input for the DSS to derive process transformation recommendations and to choose among alternative competing standards. We formulate MR2a:

MR2b: The DSS needs information concerning process standardization goals.

A challenge with transformation goals however in addition to their mutual incompatibility are different levels of importance allocated to transformation goals. Decision-making concerning process standardization thus requires multiple criteria decision-making, which requires to weigh these criteria in advance. Thus, decision-making concerning process transformation goals requires the DSS to weigh goals according to importance in advance to give one goal priority over another via a priority ranking among standardization goals [31]. We formulate MR2c as follows:

MR2c: The DSS needs an importance ranking among goals to select among alternative standard process designs.

Further, decision-making requires both forms of process knowledge to complement each other to overcome mutual weaknesses. As a direct requirement of MR1 and MR2, both types of process knowledge need to be combined in a single comprehensive as-is process model before decision-making. As bottom-up process models from MR1 and

top-down process models in MR2 each deliver an incomplete analysis of processes in isolation, both models need to be merged. Thus, we derive MR3:

MR3: The DSS needs combine both bottom-up and top-down process information in a comprehensive as-is process model for decision-making.

In addition to these status quo-oriented meta-requirements to derive a comprehensive As-Is process model, an additional meta-requirement is established concerning the possible future process state against which the as-is process model is to be matched to derive a future standard process recommendation. To select the most suitable standard process, the DSS needs to possess a repository of potential standard specifications concerning the future target process design from which an optimal process design in X' can be chosen. We formulate MR4 accordingly:

MR4: The DSS needs a repository of best-practice standard process models.

Furthermore, the purpose of the DSS is to provide decision support between the process model alternatives in the form of process standardization. One method to compare process models with standard best-practice models of ERP systems is the application of business process similarity [22] as motivated in the conceptual foundations in section 2. Thus, the proposed DSS relies on business process similarity to minimize the distance between the input models and the target models [32] in the repository (MR4). We formulate the final MR5 accordingly:

MR5: The DSS needs a similarity-based matching logic to propose an appropriate future standard process design

Following the introduction of the meta-requirements, the next section will introduce design principles to conceptualize the DSS. Section 6 contains the description of the implementation in the Apromore process analytics platform [12].

5 Design Principles and Design Decisions

We translate the meta-requirements into design principles (DPs) and design decisions (DDs) to steer the later development of the software artifact and to modularize the components of the DSS. According to MR1, the DSS is required to incorporate bottom-up process information. In turn, this requires to extract relevant process data from information systems and prepare the information for process mining in an event log database. Further, the event log needs to be visualized in a graphical process model such as a BPMN representation. Thus, DP1 is formulated as follows:

DP1: The DSS provides a bottom-up process mining layer to retrieve business processes and associated information from organizational information systems.

To account for DP1, we a data extraction program needs to be implemented in the information systems to extract the relevant process mining data (DD1.1). Further, the raw data needs to be transformed into a process mining event log (DD1.2). Finally, the process mining event log needs to be visualized graphically in a process model formalization such as BPMN to be able to unify both bottom-up and top-down information (MR3) and to perform the attribute-based similarity matching of the as-is model against the to-be standard process models. Thus, the DSS includes a BPMN visualization engine (DD1.3).

Further, MR2 requires the DSS to incorporate de jure process knowledge into decision-making, which requires the provision of a user interface to enrich the bottom-up

process mining models with additional top-down information which can otherwise not be retrieved by process mining such as additional shadow-process steps or intangible contextual process attributes outside of information systems. We formulate DP2 as:

DP2: The DSS provides the ability to enter additional top-down information.

To identify the contextual information which needs to be attached, we consulted literature on business process standardization to identify relevant process standardization attributes. In particular, the contribution by Romero et al. [33] retrieves a collection of contextual factors which impact the extent of process standardization. In their contribution, the authors find the extent of standardization to be determined by six process categories, namely process activities, resources, data, control-flow, information technology, and management. For each of these categories of contextual factors, we retrieved several sub-attributes from literature which can be assigned with either a numeric attribute value or string of characters for matching (DD2a). Thus, process models need to be attached with these top-down process standardization attributes to perform later similarity matching of the as-is process against the possible to-be standard processes. We formulate DP2a accordingly:

DP2a: The DSS provides process standardization attributes as one element of top-down information.

As standardization attributes refer to different aspects of processes such as the entire process, a specific process variant, or to the task-level, attributes need to be added to the respective level accordingly.

Furthermore, MR2b demands the incorporation of process transformation goals. Therefore, we formulate DP2b to require the DSS to provide a list of possible transformation goals. DP2a is expressed accordingly:

DP2a: The DSS provides process transformation goals as one element of top-down information.

To translate DP2a into a design decision, we performed a series of workshops with the six senior managers responsible for process transformation at the industry partner in the SAP S/4 HANA migration project to retrieve a collection of process transformation goals. Results for process transformation goals in addition to standardization include flexibility, efficiency, cost reductions, compliance, integration, process stability, transparency, measurability, simplification and complexity reductions, and sustainability, which will be given as possible matching values (DD2b).

In addition, MR2c requires the possibility to specify a relative importance prioritization to these standardization goals. Hence, DP2c demands a prioritization of the standardization goals and attributes:

DP2c: The DSS provides a priority ranking for process standardization attributes and standardization goals as one element of top-down information.

To account for DP2c, the DSS allows to weigh each attribute and transformation goal with an importance factor between 0 and 1 to adjust the relative weight assigned to the respective element in the similarity matching algorithm (DD2c).

In addition to the incorporation of bottom-up (MR1) and top-down (MR2) process information, MR3 requires to combine both types of process knowledge before decision-making in the algorithm to determine the most suited standard process.

DP3: The DSS needs to combine bottom-up process mining models and associated top-down information in an enriched process model of the as-is process.

The proposed DSS accounts for DP3 with a visualization module which combines bottom-up process mining models with standardization attributes in an enriched BPMN 2.0 model of the as-is process (X) (DD3).

To be able to propose a suited standard process specification, the enriched as-is process model needs to be matched against the different possible process designs as required by MR4. To implement the requirement, DP4 is formulated accordingly:

DP4: The DSS needs access to a repository of different best-practice standard processes designs.

To be able to perform the attribute-based similarity matching algorithm which uses the extended BPMN model of the as-is process as input, the standard process models in the repository need to be in the same format and be attached with additional top-down information. Thus, the proposed DSS contains a repository of BPMN 2.0 process models (X_s), such that the as-is process can be matched against each of the process models in the repository to determine the models with a high degree of similarity as candidate for standardization (DD4).

Finally, the last requirement MR5 refers to the need of a matching algorithm which determines the similarity of the as-is process (X) for each of the candidate process models (X_s) in the standard process repository to recommend a target model for implementation. We formulate DP5 as follows:

DP5: The DSS needs a similarity-based matching algorithm for matching of the enriched as-is process model against best-practice standard models in the repository.

Recently, “process similarity” has gained a high degree of attention and numerous approaches to process matching have been proposed. By means of a literature review, several potential process matching techniques were identified and compared to select attribute-based similarity matching as a suited candidate to solve the problem at hand. The contribution by Becker and Laue [34] categorizes process similarity measures into approaches including the correspondence between process model nodes and edges, the edit distance between graphs, causal dependencies between the different activities, and similarity approaches based on trace sets. For example, the contribution by Dijkman et al. [35] identifies five similarity dimensions to be taken into account, namely syntactic, semantic, attribute-based, type-based and contextual similarity. Therefore, the authors propose to measure the similarity from three aspects including node-matching, structural, and behavioral similarity. Finally, Thaler et al. [36] introduce natural language, graph structure, behavior, and human estimation as determinants of model similarity.

Most of similarity matching techniques are based on the model structure or behavior and define distance metrics between a pair of process models to quantify the similarity. The authors in Li et al. [37] provide an approach to measure the structural similarity between business processes based on the number of transformation operations such as adding, deleting or moving to change the structure from one business process to the other. A frequent challenge in process matching are differing labeling styles between process models. For example, a verb-object label like “create order” refers semantically to the same task as the action-noun style “creation of order”. To address the issue, the algorithm relies on natural language processing. Thus, the “BPMNDiffViz” by Ivanov et al. [38] compares process models in BPMN 2.0 language using label matching and structural matching metrics. The ICoP Framework by Weidlich et al. uses structural similarity to identify matches and correspondences between business processes [39]. In sum, the calculation of process model similarity needs to take into account

heterogeneity of behavioral representation, labeling styles and terminology [40], as well as process model structure [35]. However, for the proposed DSS, the measurement of similarity needs to be extended to take into account process model attributes such as the attached standardization information. Thus, standard process recommendations are derived through an attribute-based similarity matching algorithm which calculates process model similarity for each variant of the as-is process model against the to-be standard process models in the repository based on process model attributes, behavior, structure, and text processing of labels (DD5). Fig. 1 summarizes the conceptualization of the different modules of the DSS based on the meta-requirements and design principles.

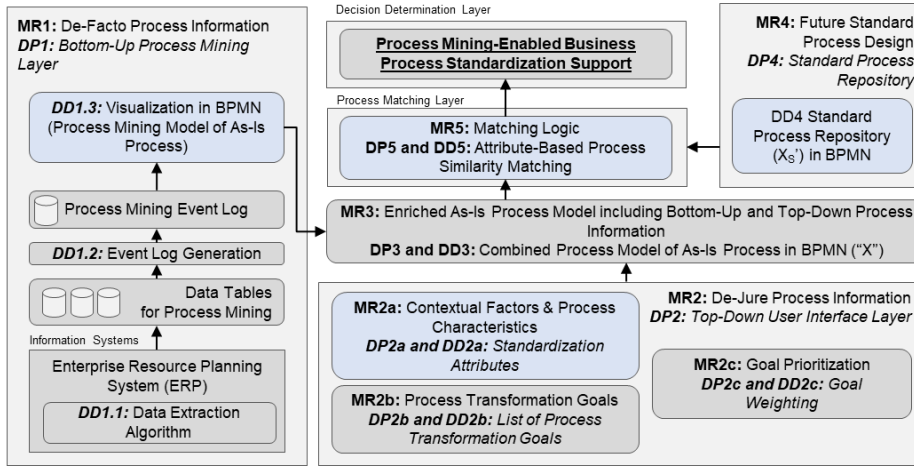


Fig. 1. Conceptualization of the Decision Support System and Modules (Blue: Essential)

6 Implementation in the Apromore Platform

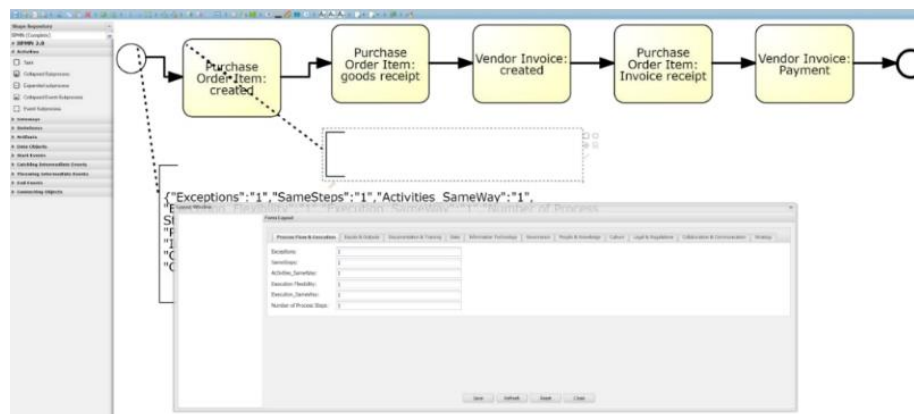
We associate a number of benefits to the implementation in Apromore, which is an open-source collaborative online business process analytics platform provided by the Apromore Initiative [12]. Specifically, with regard to both the wide acceptance in the community and the rich functionalities provided, we decided to implement the DSS in Apromore. In addition to the workshops performed in the SAP S/4 HANA project context at the industry partner to enrich the meta-requirements from academia with practical insights, the Apromore DSS uses real-world data from three SAP R/3 ERP systems from three sub-companies of the manufacturing corporation.

To account for DP1, we implemented a data extraction program in each of the SAP R/3 systems of the industry partner to extract the relevant data tables required for process mining as .csv-files (DD1.1). Further, the raw data in individual .csv files needs to be translated into a process mining event log. Thus, our solution imports all relevant data into an SQL database to perform the event log generation by a SQL transformation script. To perform the event log generation, a German process mining company provided the transformation script for the purpose of this research to generate the event log from the SAP raw data (DD1.2) in the SQL database. Finally, we export relevant information from the event log into .xes-files for Apromore.

Table 1. Overview over Process Mining Event Log

Process	Purchasing			Sales		
Company	A	B	C	A	B	C
Period	01.01.2016 - 31.07.2017					
Number of cases	998,80 Thsd.	432,21 Thsd.	108,54 Thsd.	15,8 Mil.	65.377	155.125
Number of process variants	20,67 Thsd	10,47 Thsd.	2,54 Thsd.	35,32 Thsd.	39,815 Thsd,	20,87 Thsd.
Total number of process steps [Millions]	4,13	2,15	0,34774.	106,52	50,49	6,07
Avg. number of process steps	4,13	4,98	4,42	6,74	6,02	8,37
Distinct process steps	30	154	54	21	21	22

In principle, the DSS implementation in Apomore can be adapted to incorporate other and any forms of process mining event logs. Finally, we use the BPMN visualization functionality provided by the Apomore platform [12] to draw process models (DD1.3).

**Fig. 2.** Graphical User Interface to Attach Top-Down Process Information

Further, the Apomore instantiation provides a graphical user interface as illustrated in Fig. 2. The user interface allows to attach standardization attributes which are valid for either the entire process, a particular process variant, or a specific task (DD2a). Further, project information such as transformation goals can be entered through a list of possible transformation goals (DD2b) and be prioritized through a numeric weight factor between 0 and 1 (DD2c).

Furthermore, the information needs to be combined in an enriched BPMN process model of the as-is process (X) according to (DD3). We use BPMN annotations for visualization in Apomore to display the additional top-down process information.

To retrieve the repository of standard business processes according to DD4, we downloaded the library of standard process specifications from the SAP SolutionManager and imported the library into Apomore as matching candidates. In addition to other ERP management functionalities, SAP SolutionManager is a comprehensive tool to perform business process management and documentation for the SAP ERP

landscape. SAP SolutionManager 7.2 provides a publicly available database of to-be standard processes in BPMN 2.0 language for SAP systems. Each of the to-be process models was enriched with the standardization attributes and assigned with values in a workshop with 6 process experts to be able to implement the algorithm.

Finally, to perform the attribute-based similarity matching according to DD5 under consideration of the additional standardization attributes and process transformation goals, we developed a new similarity-based matching plugin based on the existing “similarity search” plugin in Apomore. The algorithm performs matching in three steps. The first-level matcher performs matching at the process-level to ensure the as-is process is matched against the correct domain of the to-be processes such as sales or procurement processes in the repository and considers process-level standardization attributes. Further, each variant of the as-is process differs from the other variants in terms of graph structure, variant behavior, and standardization attributes. Thus, the second variant-level matcher calculates the similarity of each variant of the as-is process according to behavior, graph structure of the variant, and the difference between attribute values. Third, the task-level matches similarity of tasks and attributes. Compared to existing approaches, process models do not contain additional top-down information such as process standardization attributes, which requires the algorithm to consider similarity of attached standardization attributes. For each top-down attribute, the numeric distance is computed. Distances are multiplied by attribute weights and divided by the number of attributes to achieve a weighted similarity score of an individual task within the variant. The overall similarity for a to-be process in the repository is calculated by the sum of variant similarities weighted by the number of variant occurrences. The final result of the attribute-based similarity-matching algorithm in the DSS is thus a similarity measure between 0 and 1 (1= perfect similarity) for each of the to-be standard processes in the repository. Thus, decision-makers receive a list of all standard processes ordered by descending similarity to the as-is process. The algorithm displays the final similarity score report for each of the to-be standard process models in the repository ordered by descending similarity.

7 Evaluation

Evaluation of the artifact quality is a critical element of any DSR project (e.g., [29]). To determine the ability of the DSS artifact to achieve the intended purpose, three aspects of the DSS are evaluated separately in the first design cycle. First, feasibility of the DSS is evaluated by applying the DSS for the purchasing and the sales process in company A. Second, process decision-makers are asked for their experiences on the DSS in semi-structured interviews. Third, visualization forms of the attribute-extended process models are evaluated in terms of process comprehension. In the second design cycle, an evaluation of the process matching algorithm will be conducted by means of comparison against the matching performed by human users.

In the technical feasibility evaluation, we considered the number of variants to cover a threshold of at least 80% of cases for each process. For the purchasing process of company A, 41 variants were taken into account which cover a span of 869,63 thousand purchase orders and assigned with the standardization attributes on the process-, variant-, and task-level in a workshop with three purchasing process experts. After

application of the similarity matching algorithm, the proposed target standard process was the standard end-to-end procurement process from SAP which achieved the highest similarity score of 0,87. Likewise, for the sales process of the company, 56 variants were processed to cover 12,74 million sales orders. As the as-is process contains a large number of customer-specific adaptations, the algorithm produced a comparably low degree of similarity of 0,68 for the SAP standard process specification “Sales from Stock Direct Sales” for the new S/4 HANA ERP system. Table 2 presents results for the application of the DSS for the purchase-to-pay and the order-to-cash processes for one sub-company of the manufacturing corporation.

Table 2. DSS Results for Purchasing and Sales Process of Company A

Process	Purchase-to-Pay ("Purchasing")	Order-to-Cash ("Sales")
Company	A	A
Number of cases considered	869,63 Thsd.	12,74 Mil.
Number of variants considered	41	56
Number of different tasks	30	15
Similarity score of proposed standard process	0,87	0,68
Proposed target standard	SAP_E2E_P2P Standard_Procurement	SAP_E2E_O2C Sales_from_stock_Direct_Sales

When asked for their opinion on the DSS and the helpfulness in process standardization, managers highlighted the ability of the DSS to support the selection of a suitable standard process and to justify the decision due to the reliance on data from the ERP system from process mining. Further, managers liked the DSS as it allows for analyses of the required changes to the process before the implementation of the new standard process. Managers further stated the DSS further helped them in advancing BPM as a core capability of the organization, and to increase the “process-oriented thinking” of their employees and themselves. However, managers further highlighted the effort to attach all top-down standardization information to the process variants, and the requirement to implement process mining in a pre-project.

Second, four different forms of representation of the extended process models were evaluated in terms of process model comprehension to determine how well process decision-makers are supported in their understanding of processes and the associated standardization information. A structured literature review was undertaken to identify impact factors on process model comprehension. Based on the results, three influencing factors on comprehension, namely visualization, decomposition and visual guidance were selected for the design of enriched process models which differ in their representation of the standardization attributes. The first representation lists the standardization attributes in an additional table next to the BPMN process model. The second visualization statically integrates the attributes directly into the BPMN process model and links the attributes in branches to the entire process, the variant, or an individual task, respectively. The third visualization form copies the second form, but additionally provides an interactive component which lets users dynamically show and hide the attributes. The fourth process model realizes visual guidance features and illustrates attributes and values with icons. These four process models were evaluated by an online experiment with $n = 8$ process experts and $n = 35$ students regarding their impact on the dependent variable process model comprehension in terms of the time required to

answer comprehension questions to the process models and the standardization attributes. Comprehension was operationalized by effectiveness (the number of correct answers), efficiency (time spent on answering) and relative efficiency (proportion of effectiveness through efficiency) as well as by the subjective measurement of perceived ease of comprehension. Results indicate that the static visualization form without the dynamic interactive feature achieves the highest process model comprehension in terms of efficiency and relative efficiency. Besides, the fourth visualization form with visual guidance has the highest effectiveness and is perceived as the easiest form to comprehend. Regarding subjective preferences of respondents, respondents preferred the guided process model with icons ($n = 20$ respondents who ranked the variant as their highest preference; 46,51%) and the static process models extended with branches ($n = 12$; 27,91%), while the interactive process model extended with branches ($n = 4$; 9,30%) and the process model extended with a table ($n = 1$; 2,33%) achieved the lowest result. Thus, the second design cycle will implement the guided process model with icons.

8 Conclusion and Outlook

Although state-of-the-art information systems increasingly provide organizations with tremendous amounts of process data, and process mining delivers mature techniques to turn data into process information, turning information into actual process decisions remains a substantial challenge. This paper designs a process mining-enabled DSS to aid organizations in process standardization. By extending “de-facto” process models from process mining with additional “de jure” process information in decision-making, the DSS might considerably improve the ability to standardize business processes. However, the DSS also encounters several limitations and requirements to the second design cycle. First, the DSS determines the process model with the highest degree of similarity from the repository of best-practice standard processes. Although “similarity” implies a minimization of organizational change and thus lower tangible and intangible costs for implementation, the “best” candidate for implementation might be a more radical change towards a process with only a low degree of similarity to the as-is process. Second, to match business process against models in the repository, the to-be standard models need to be attached with top-down information, which might differ between organizations and thus not generalize to other contexts.

References

1. Teece DJ (2010) Business Models, Business Strategy and Innovation. *Long Range Planning* 43(2-3): 172–194
2. Tiwari A, Turner CJ, Majeed B (2008) A review of business process mining: State-of-the-art and future trends. *Business Process Management Journal* 14(1): 5–22
3. van der Aalst WMP, Weijters AJMM (2004) Process mining: a research agenda. *Process / Workflow Mining* 53(3): 231–244
4. van der Aalst WMP (2011) *Process mining: Discovery, conformance and enhancement of business processes*. Springer, Berlin, Heidelberg, New York

5. van der Aalst WMP (2014) Process Mining in the Large: A Tutorial. In: Zimányi E (ed) Business Intelligence, vol 172. Springer International Publishing, Cham, pp 33–76
6. Indulska M, Green P, Recker J et al. (2009) Business process modeling: Perceived benefits. A.H.F. Laender et al. (Eds.): ER 2009, LNCS 5829: 458–471
7. van der Aalst WMP (2013) "Mine your own business": using process mining to turn big data into real value. Proceedings of the 21st European Conference on Information Systems
8. Loebbecke C, Picot A (2015) Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems* 24(3): 149–157
9. van der Aalst WMP, Reijers HA, Weijters AJMM et al. (2007) Business process mining: An industrial application. *Information Systems* 32(5): 713–732
10. Wurm B, Schmiedel T, Mendling J et al. (2018 (forthcoming)) Development of a Measurement Scale for Business Process Standardization. ECIS 2018 Research-In-Progress Papers
11. Botta-Genoulaz V, Millet PA, Grabot B (2005) A survey on the recent research literature on ERP systems. *Computers in Industry* 56(6): 510–522
12. The Apromore Initiative (2018) Apromore: Advanced Process Analytics Platform. <http://apromore.org/about>
13. Donaldson L (2001) *The Contingency Theory of Organizations*. Sage
14. Sousa R, Voss CA (2008) Contingency research in operations management practices. *Journal of Operations Management* 26(6): 697–713
15. Van Der Aalst, Wil MP, Dustdar S (2012) Process mining put into context. *IEEE Internet Computing* 16(1): 82–86
16. Melão N, Pidd M (2000) A conceptual framework for understanding business processes and business process modelling. *Information systems journal* 10(2): 105–129
17. Vom Brocke J, Zelt S, Schmiedel T (2016) On the role of context in business process management. *International Journal of Information Management* 36(3): 486–495
18. Škrinjar R, Trkman P (2013) Increasing process orientation with business process management: Critical practices'. *International Journal of Information Management* 33(1): 48–60
19. Galbraith JR (1973) *Designing Complex Organizations*. Reading: Addison-Wesley
20. Haußmann C, Dwivedi YK, Venkitachalam K et al. (2012) A Summary and Review of Galbraith's Organizational Information Processing Theory. In: Dwivedi YK, Wade MR, Schneberger SL (eds) *Information Systems Theory*, vol 29. Springer New York, New York, NY, pp 71–93
21. Sprague RH (1980) A Framework for the Development of Decision Support Systems. *Management Information Systems Quarterly* 4(4): 1
22. Fischer M, Heim D, Janiesch C et al. (2017) Assessing Process Fit in ERP Implementation Projects: A Methodological Approach. *International Conference on Design Science Research in Information Systems 2017*, Karlsruhe, Germany, May 30 - June 1, 2017
23. Harmon P (2010) *Business process change: A guide for business managers and BPM and Six Sigma professionals*. Morgan Kaufmann

24. Nissinboim N, Naveh E (2018) Process standardization and error reduction: A revisit from a choice approach. *Safety Science* 103: 43–50
25. Schönig S, Cabanillas C, Jablonski S et al. (2016) A framework for efficiently mining the organisational perspective of business processes. *Decision Support Systems* 89: 87–97
26. Agrawal R, Gunopulos D, Leymann F (1998) Mining process models from workflow logs. *Computer Science*. IBM, San José, California
27. Markus ML (2004) Technochange management: Using IT to drive organizational change. *Journal of Information Technology* 19(1): 4–20
28. Hevner AR, March ST, Park J et al. (2004) Design Science in Information Systems Research. *Management Information Systems Quarterly* 28(1): 75–105
29. Venable J, Pries-Heje J, Baskerville R (2016) FEDS: a Framework for Evaluation in Design Science Research. *European Journal of Information Systems* 25(1): 77–89
30. Rosemann M, Vessey I (2008) Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks. *Management Information Systems Quarterly* 32(1): 1–22
31. Afflerbach P, Leonhard F (2016) Customer Experience Versus Process Efficiency: Towards an Analytical Framework About Ambidextrous BPM: Completed Research Paper. *The Journal of Strategic Information Systems*
32. Martens A, Fettke P, Loos P (2014) A genetic algorithm for the inductive derivation of reference models using minimal graph-edit distance applied to real-world business process data. *Tagungsband Multikonferenz Wirtschaftsinformatik*: 1613–1626
33. Romero HL, Dijkman RM, Grefen PWPJ et al. (2015) Factors that Determine the Extent of Business Process Standardization and the Subsequent Effect on Business Performance. *Business & Information Systems Engineering* 57(4): 261–270
34. Becker M, Laue R (2012) A comparative survey of business process similarity measures. *Process / Workflow Mining* 63(2): 148–167
35. Dijkman RM, Dumas M, van Dongen BF et al. (2011) Similarity of business process models: Metrics and evaluation. *Information Systems* 36(2): 498–516
36. Thaler T, Schoknecht A, Fettke P et al. A Comparative Analysis of Business Process Model Similarity Measures. *BPM 2016 Workshops*: 310–322
37. Li C, Reichert M, Wombacher A (2008) On Measuring Process Model Similarity Based on High-Level Change Operations: Conceptual Modeling - ER 2008. *International Conference on Conceptual Modeling*
38. Ivanov SY, Kalenkova AA, van der Aalst WMP (2015) BPMNDiffViz : a tool for BPMN models comparison. *Business Process Management Conference 2015 Demo Papers*, Innsbruck, Austria, August 31 - September 3, 2015
39. Weidlich M, Dijkman RM, Mendling J (2010) The ICoP Framework: Identification of correspondences between process models. *International Conference on Advanced Information Systems Engineering*: 483–498
40. Dijkman RM, van Dongen BF, Dumas M et al. (2013) A Short Survey on Process Model Similarity: Seminal Contributions to Information Systems Engineering: 25 Years of CAiSE: 421–427