



A systematic approach for business service consolidation in virtual organizations

Hodjat Soleimani Malekan¹ · Katarzyna Adamiak¹ · Hamideh Afsarmanesh¹

Received: 16 February 2017 / Revised: 13 July 2017 / Accepted: 20 December 2017 / Published online: 8 January 2018
© The Author(s) 2018. This article is an open access publication

Abstract

Various forms of partnerships have increasingly emerged among organizations, in both the manufacturing and service industry. In one type of structured partnership, called the Virtual Organization (VO), the member organizations share with each other a part of their capacities and capabilities, so that together they can seize a bigger market share and more opportunities. However, to excel the VO's potential, the participating organizations must become fully informed about each other's particular abilities and competencies, as represented by their shared business services (BSs). Our approach proposes the creation of a common shared pool of integrated BSs, and further optimizing it through consolidating similar and related BSs, to reduce the size of this shared pool and enhance its learning curve by the VO partners. To this end, we extend and adapt a Business Process merge technique to make it more suitable to BS consolidation and VO applicability. In this paper, a semiautomated methodology is introduced that creates an optimized, clean (i.e., anomaly-free consolidated model), and valid (i.e., approved by domain experts) set of consolidated BSs in the VO context. Design science principles are applied to this research, and a systematic literature review is also performed. Furthermore, a real case study from the construction industry is conducted to validate the introduced methodology.

Keywords Virtual organizations · Business service consolidation · Business processes

1 Introduction

In a large number of industries, organizations are forming partnerships to provide business services that they could not provide individually. One type of structured partnership is called the Virtual Organization (VO) that represents a dynamic network within which independent organizations share a part of their abilities and resources [1] in order to target more market opportunities, create innovative products, improve quality and reduce costs, among others [2]. In parallel, organizations increasingly formalize definition of their business services, to both benefit service implementation internal to their organization, as well as for advertising their

competencies to the external world [3]. This in turn motivates the VO partners with relying on the Service-Oriented Architecture approach, not only to facilitate creation and provision of value-added BSs [4], but also to gain advantage from the possibility of online execution of the BSs [5].

In this research a semiautomated methodology is first introduced and then applied to a real-world case of VO partnership in the construction industry. Partnership formation is traditionally a popular practice in construction industry, however, often performed in unstructured way, where different organizations are specialized in different discipline and individually fulfill multiple tasks [6]. While some tasks are very different, others are common to all, such as the price negotiation with the material providers and the procurement.

Similarly, in our real scenario, the organizations used to co-work during the life time of every such project with no need to share and/or merge their BSs with each other. Therefore, each involved organization executed individual BSs during different stages of the joint project, and only interacted with the BSs provided by others. However, lack of sharing of BSs prohibited them from working more efficiently together and delivering composed value-added BSs [7]. Moreover, in

✉ Hodjat Soleimani Malekan
h.soleimanimalekan@uva.nl

Katarzyna Adamiak
adamiakkasia@yahoo.com

Hamideh Afsarmanesh
h.afsarmanesh@uva.nl

¹ Informatics Institute, University of Amsterdam,
P.O. Box 94323, 1090 GH Amsterdam, The Netherlands

case of every new joint construction project, same partners again needed to spend time and efforts on establishing similar BS interaction relations from one partner to another.

Establishing a VO in a niche market in construction industry is exemplified in this paper, where it targets construction of eco-friendly buildings. In 2016, it is forecasted that the number of interested investors in this type of projects will increase by about 60 percent till 2018 [6]. However, construction of eco-friendly buildings requires the use of novel technology, materials, and expertise from various disciplines. Through effective cooperation among a set of organizations involved in a VO, they can successfully deliver such niche projects.

Continued collaboration among organizations and sharing all their abilities in the VOs create the base for partners to provide more efficient wider range of services and shorten the time-span of operation. While in the past, organizations who worked on joint projects were reluctant to share and/or consolidate their BSs, today once a VO collaboration contract is negotiated, organizations are more willing to do so; especially once realizing the advantages resulted from BS consolidation in VOs [4,7].

Our approach to BS consolidation assists organizations to concisely define their own practiced services as well as using each other's experiences and best practices, therefore underpinning potential efficient collaboration in future. Another advantage targeted by this paper is that the BS consolidation facilitates formation of an optimized and consistent shared pool of consolidated BSs at the VO level. Optimization would involve for instance, identifying different variations of the same BS, shared by different VO partners, and unifying them into one BS. Therefore, further to enhancing common understanding about existing BSs by all VO member organizations, it facilitates the creation of composed and value-added BSs. Our BS consolidation method targets creating one BS model out of several smaller BS models. However, there is a lack of research addressing consolidation of formalized BSs. Nevertheless, since every BS comprises a set of business processes (BPs), we fall back on investigating BP consolidation approaches, as also represented in this paper, and adapt a BP merge technique to elaborate our consolidation approach in the VO context [5].

The *design science* principles are applied to our research for the design of the BS consolidation methodology. As a part of these principles, to build reliable research background against the requirements for our research, a systematic literature review (SLR) [8] is first performed, identifying research gaps in the area. We illustrate the results of our SLR and the gaps in essential stages for consolidating several BS models. It is shown that step-wise definition of a balanced BS consolidation methodology is missing in approaches introduced in related research, such as in [9,10], and not empirically validated in a VO context. Moreover, there the existing

semiautomation approaches for BS consolidation do not efficiently support the users, due to the complexity of their produced results [5]. Then, we introduce a semiautomated methodology needed to constitute the optimized set of processes in the VO context and evaluate its appropriateness in a case study [11].

To fully understand and address the above research concerns, three research questions are discussed:

- RQ.1 What is the state of the art in BS/BP consolidation through applying merge techniques, and what are the main stages thereof?
- RQ.2 Which particular steps are needed to be taken by VO member organizations toward creating a common set of consolidated BSs?
- RQ.3 How to ensure the introduced methodology is valid and applicable to the VO context?

This paper is structured as follows. Section 2 introduces the scientific framework applied to our performed research. Section 3 addresses our finding from SLR. Our BS consolidation methodology is introduced mainly in Sect. 4. The suitability of our introduced methodology for an example case from construction industry is represented in Sect. 5. Finally, the paper is concluded and discussed in Sect. 6.

2 Research approach: design science

The information systems research (ISR) framework defined for *design science* facilitates study of a phenomenon to design an information system. This framework, defined by Hevner et al. [12], is rooted in problem-solving paradigm.

In principle, ISR focuses on building and evaluating designed artifacts, such as models and methods. These artifacts are utilized to improve the efficiency and effectiveness of organizations. In order to build a novel artifact, the ISR framework [12] associates the environment of the information system (e.g., VO context) to the knowledge base of related theories and methods, e.g., for BS consolidation. Three different research cycles, including *relevance*, *rigor*, and *design* are also applied to connect the design science context to the environment space and knowledge base space, as shown in Fig. 1.

The purpose of our presented artifact is to contribute to the development of the field by introducing a novel BS consolidation methodology for the organizations cooperating within a VO. It is important to ensure that the presented solution is relevant and useful in practice. Therefore, our methodology artifact is created by following and complying with seven guideline criteria for proper design science research, as introduced in [8] and briefly explained below.

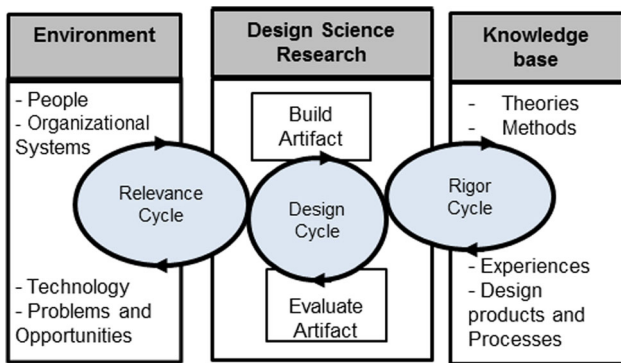


Fig. 1 ISR framework as explained in [12]

1. *Design as an artifact*—the methodology of BS consolidation is a viable artifact, aimed at providing practical support to created VOs.
2. *Problem relevance*—the artifact supports important and relevant business problems (e.g., redundant BP models). VO formation is currently an increasingly significant and challenging practice, counteracting market turbulence and striving to ensure survivability. This is especially important for SMEs, which are the core of interest in our research, as stated in [2].
3. *Design evaluation*—our methodology is evaluated based on a real case study of a VO, consisting of ten individual organizations that join efforts to increase their operational profitability and jointly provide effective business services. The methodology steps are applied to this case on one hand to examine usability and on the other hand to highlight limitations for future improvement. This evaluation is conducted through a case study in accordance with [11].
4. *Research contributions*—contributions of the proposed artifact include review of current state of the art in the VO context and of relevant Business Process Management principles, as well as systematic literature review [8] of the existing theories and methods for BP consolidation. Moreover, conducting the case study significantly supports applicability of our methodology in consolidating BS/BP models to support VOs.
5. *Research rigor*—rigorous research methods are applied to both: identifying the gaps and building the research foundations through the SLR, and evaluating every step of the proposed artifact through the real case study.
6. *Design as a search process*—addressing the identified gaps, each methodology step both points to certain identified gap, and is designed by applying the state of the art in research when it addresses the needs for that step, e.g., the needed algorithms were developed for the BP graph similarity evaluation. Thus, it assures searching and identifying valuable existing alternatives, as well as selecting reliable methods from the existing knowledge base.

7. *Communication of research*—we aimed at presenting the methodology that is understood and usable by both: management-oriented and technology-oriented audiences. It can be used directly by organizations interested in integrating their BSs within a created VO, as well as by software engineers developing software for automated BS consolidation in VOs. In summary, the discussed guidelines highlight that our proposed methodology conforms to design science standard. Furthermore, our BP consolidation methodology addresses both the business needs of VO context and applies the knowledge in BP management discipline, sufficiently and rigorously.

3 State of the art in BP consolidation

To identify the state of the art in BP consolidation based on the merged technique, a systematic literature review is performed, and hence two main search questions are defined in accordance with steps explained in [8]. Then, our findings are analyzed to identify the existing gaps. Finally, the aimed contributions are discussed and contrasted against main criteria.

3.1 Systematic literature review

Understanding a rigorous procedure, SLR provides unbiased studies of the related publications, through the following steps: i) justifying the review background and setting review questions, ii) defining the search protocol of the review, including “specifying data sources” (by reviewing title and abstracts), iii) selecting primary studies based on data inclusion and exclusion criteria, iv) finalizing findings from the primary studies based on appropriate assumptions to ensure aligning with the search protocol (by regarding full text of papers), and v) discussing the validating the search process from various aspects, e.g., highly cited papers, approach most prolific researchers.

We mainly address through SLR two parts of our RQ.1 question in this paper (See Sect. 1) in two separate search questions on: (i) existing BP consolidation approaches, which are supported by merging techniques, and (ii) the extent that consolidation of BP models requires further research based on its stages. With the latter issue, we search for potential improvement lines for BP consolidation methodology. For the SLR purpose, Google Scholar is used. Primary concern in searching scientific resources is to determine the searching queries. The term “consolidation”, as defined in Sect. 1, is related to different keywords in BP management discipline, while it is new in relation to BS management in VOs. In principle, consolidation shares several ideas and concepts with BP management approaches that imply the combination of different BP models into

one artifact in a certain procedure. The merge technique is mainly used in BP consolidation procedure to create a merged model [10].

To develop a merged BP model, various parts of different BP models are indistinguishably combined into a completely new BP model, e.g., in [13,14]. Sometimes, a family of BP variants are merged into a so-called configurable merged model [5], where a BP model could be further customized, e.g., [10,15]. Studies have suggested that the closest concepts to consolidation are integration, variability, and merging. However, integration and variability aim at exchanging information and messages without unifying the BPs, whereas our consolidation concerns with finding common parts of the models and resolving differences between them, through merging models. Thus to conduct our search, different grammatical forms of the two words “merge” (i.e., “merging” and “merged”) and “consolidation” (i.e., “consolidate”, “consolidated”, and “consolidating”) are combined with the “business process /service” (or business processes/services) and process model (or Process models).

For each query, we select the first 200 hits of the Google scholar, from 2000 to 2015 and filter them based on the titles. We consider journal papers, conference papers, and book chapters. However, we exclude duplicates as well as less than 5-page papers, and papers without citations. Furthermore, we give particular importance to cited articles and forward-check the papers that cite it, and backward-check the papers in the references thereof up to first 50 hits. During the search, we find out that the concept of “aggregation” (in different grammatical forms, such as “aggregation”, “aggregate”, and “aggregating”) is also related to the consolidation approaches. Besides, workflow is used instead of business process interchangeably. Therefore, we search other queries by considering new key words (i.e., “workflow” and “workflows”). After checking the abstracts of all search results, 86 papers were selected as the results of the primary search.

To verify the preciseness of our primary search, we identified four related literature surveys. First, Ayora et al. [16], review the variability of BP models in process-aware information systems. The survey considers related concepts to variability management of BP modes such as configuration, merging, and evolution. From [16] we select the merge supporting our approaches. The other survey is by La Rosa et al. [17] about business process variability modeling. From this one, we pick the consolidation supporting approaches. Moreover, van der Aalst in [5] and Dijkman in [18] introduce some merge-based consolidation approaches. In all cases, our primary search results subsume and extend the findings of all these four surveys, in BP merge aspects. Moreover, we follow the references of these studies, to ensure that nothing is left out. After reading the full text of papers, we identify a list of ten papers, which address the BP consolidation based on merge techniques, constituting the state of the art related to

our BS consolidation research. The summarization and data extraction step is performed in two steps including general introduction of finding, and systematic evaluation of consolidation approaches. The evaluation is performed by two first authors, and the results, including disagreements, are further discussed in a conflict resolution meeting.

Ten papers present methods, frameworks, or approaches to consolidate BP mode by merging them. Below, a summary of each paper is provided. Note that also each of these papers constitute one row in Table 1.

1. Regarding two scenarios of organizational merge and BP re-engineering, Sun et al. [19] introduce a method for merging block-oriented workflows. Given two input process models, first a merged model from their common tasks is built. Then, the variants of input models are added by four categories of merge operations “*sequential*”, “*parallel*”, “*conditional*”, and “*iterative*”. In sequential merge of two variants, either one of them is kept or both add to the merged model sequentially. By using AND gateway, two variants are inserted into the merged BP model in parallel situation. To execute one or both variant fragments of the input models based on certain conditional situation, OR gateway is used to connect them to the merged model. The iterative merge of variants occurred when the variants appear in a loop. For instance, based on preceding common tasks, several executions of the variants are needed.
2. Mendling and Simon [20] introduce a method for consolidating different views of BP stakeholders. Where, an integrated model is assumed as a database schema and opinions of different stakeholders are added as data to this schema. In this respect, the method first identifies the *semantical* relationships between constructs of two given BP models. The relationships could be either *equivalence* if two constructs are completely similar, or *sequential* if two constructs are modeled based on different views. By using AND-join gateway, the sequential constructs are added to the equivalent constructs. Finally, a set of restructuring rules eliminates some extra gateways. For example, change two subsequent direct AND gateways to one AND gateway.
3. Gottschalk et al. [9] present an approach to merge two Even-driven Process Chain (EPC) models by constructing a *function graph*. First, the two input EPC models are reduced into their corresponding function graph, where only the functions and logical control flow between each function are reserved. For example, if no connection exists, then just an XOR construct is annotated on the function graph. Then, two function graphs are combined into a merged function graph. Accordingly, an EPC model is extracted from the merged function graph.

Table 1 The SLR evaluation summary

Study	1C	2C	3C	4C	5C	6C	7C	8C	9C	10C	Comments
[19]	Ex	STL	Block-structured	Sy	N	N	Merging workflows	N	N	C, F	Not fully implemented/Less understandability because of using block-structured languages
[20]	Ib	STL	EPC	Sy	N	N	Improving process	±	±	C, F, I, V	Risk of deadlock occurrence because using AND-join/Not fully implemented
[9]	Ib	STL	EPC	Sy	N	N	Standardization	±	±	C, F, I, V	Lack of scalability to other languages (e.g., BPMN)/Not evaluated empirically
[13]	Ex	STL	IBM modeler	Sy	N	±	Standardization	±	±	C, I±	Focus is on a tool not a method/Not scalable because of focus on IBM modeling notation/not evaluated empirically
[21]	Ib, Ir	STL	aEPC	Sy	N	N	Understandability	±	N	C, I, V	Limited in preserving the correctness/modeling language limitation (EPC)
[22]	Ex	TS	Block-structured	Sy	N	±	Minimizing changes	N	N	C, F, I, V	No checking soundness of the merged model/Not include the behaviors of the input model
[23]	Ib, It, Ir	TS	EPC	Sy	N	N	Configurability	N	N	C, I, V	No framework for supporting the configurable merge
[15]	Ib, It, Ir	TS	CoSeNets	Sy, Bh	N	N	Configurability	N	±	C, F, I, V	Not addressed evolution of merged model
[10]	Ib, It, Ir	TS	EPC, BPMN	Sy, Bh	N	±	Configurability	N	N	C, F, I, V	Lack of considering end-user opinion/ Not evaluated in a merger context
[14]	Ex	STL	BPMN, BPEL	Sy, Bh	N	N	Improvement	N	±	C, F, I	Lack of considering end-user opinion/ not evaluated in a merger context
[**]	Ib	TS	BPMN	Sy, Bh, AR	Y	N	BS consolidation	Y	Y	C, I±, V	Supporting VO context

Legend: In 1C- Ib:Include behavior, Ir:Include traceability, Ir:Include reversibility, Ex:Exclude behavior/In 2C- TS:Textual Sensitive, STL:Same Textual Label/In 4C- Sy:Syntactical, Bh:Behavioral, AR: Ambiguity Resolution/In 5C-& 6C- & 8C- & 9C- Y:Yes, N:No, ±:Partially / In 10C- C:Conceptualization, F:Formalization, I:Implementation, V:Validation, ±:Partially

4. To assist modeler for consolidating BP models, Kuster et al. [13] provide a semiautomated approach in three steps. First, by comparing two given input models, differences between models are identified based on single-entry single-exit approach. Second, these differences are categorized and visualized for modeler in IBM Web-Sphere Business modeler. Finally, resolving differences by modelers preference is done, iteratively.
5. To alleviate the problem of understandability in complex big configurable merged models, Reijers et al. [21] propose a method for maintaining and retrieving model variants, during the modeling project in organizations. This approach introduces *aggregate EPC* (aEPC), which associate product in addition to existing EPC notation. The products relate the special business context to EPC diagram. The complexity of variability (e.g., conditions for different BP variants) are resolved and kept in the product elements, instead of appearing on the EPC diagram. Since the modelers need single simple and not detailed models that shows similar BPs in an organization, a (mostly manual) approach is introduced for building an aggregated BP model.
6. Li et al. propose an approach for merging BP variants into a merged model, by regarding *minimum change distances* input variants [22]. The change distance is one of basic change operations, including insert, delete, and move. Accordingly, the total operations to change each BP variant model into merged model should be minimal. The approach is adapted for block-structured languages.
7. Instead of merging two models at the same time, Derguech and Bhiri propose an approach for merging several variants of BP models [23]. The approach is tuned for EPC and relies on two preprocessing and post-processing steps for making a configurable merged model. Moreover, the approach considered single-entry single-exit structure to make merged model. Finally, the *consecutive* and *trivial* gateways such as two succeeded AND gateways are eliminated. The approach is experimented in a logistic case study.
8. Introduced by Schunselaar et al. [15], CoSeNets is a tree-like representation of block-oriented BP models, which is used for correcting the merging BP models. The CoSeNets structures of each input model are merged and any required restrictions for further extraction of input models are connected to the merged model.
9. La Rosa et al. propose an approach for consolidating two graphical BP models by merging them. For this purpose, the following tasks are performed semiautomatically. First, *common regions* are matched. Second, the *differences* in BP models are appended to common region and produce a configurable merged model. Finally, *entanglements* such as circular loops are resolved from the merged model. This approach finds the intersections

of the BP models as the potential parts for further reusing in the merge of BP models [10].

10. In a model-driven approach, Greth introduces a framework for BP model change management in [14]. It suggests merge of different versions of BP models, which could be formalized in different modeling notations (e.g., BPMN or EPC). In order to perform a language-independent merge, first an *intermediate representation* (IR) of the input model is built. Then, the corresponding elements and differences of IRs are identified and mapped together. Thereafter, the dependencies, equivalences, and conflicts of IRs are analyzed. Finally, the IRs are merged and transformation of IRs to source languages is done.

Ten criteria that are also extracted from the literature [5,16,17] in our SLR, for data extraction purposes, we evaluate the required data through . Each of the 10 papers are then evaluated against these ten criteria. The results show to what extent each paper supports the BP consolidation based on merge. The criteria are as listed and briefly described below. **1C.Preserving process behavior** : this criterion refers to the concept if the consolidated models include all the behavior of the input BP models, or if they exclude their behavior. Moreover, the traceability, i.e., finding a process in merged model, and reversibility, i.e., extracting process from the merged model, are considered as the complementary levels of behavior inclusion. **2C.Presuming similar labeling** : labeling is an important way of communicating through modeling. However, approaches deal with heterogeneity and homogeneity of labeling differently. We identify whether same labeling is assumed in the approaches or if the approaches are label sensitive. **3C.Depending on special modeling languages**: approaches can be either language-dependent or independent. For instance, certain approaches are suitable for graphical languages (e.g., BPMN or EPC), while some consolidate BPs based on an abstract model that is language-independent. **4C.Preserving merged BP model correctness**: the merged model is comprised of input models, while during the consolidation process there is a risk of errors occurring in the merged BP model. Accordingly, some approaches consider different levels of correctness, such as syntactical, i.e., concerned with structural consistency of the merged model, for instance by avoiding a disconnected node, or behavioral, i.e., by avoiding anomalies such as deadlocks or livelocks in the merged models and resolving ambiguous patterns such as OR-join. **5C.Considering collaboration in consolidation**: the consolidation process is collaboration-intensive in principle. Different stakeholders are involved in BP consolidation, either intra-organizational or inter-organizational. Therefore, the approaches are evaluated for supporting collaboration process. This criteria specifically addresses the need in VOs to establish collaboration in consolidation procedure. **6C.Preserving the**

evolution concerns: since the approaches need to focus on the entire lifecycle of the merged model, the support for evolution concerns in approaches is evaluated. **7C.Constituting goal of consolidation:** the main intention of the consolidation approach is important. Thus, identifying the main aim of consolidation is a significant factor. **8C.Including user preference:** this criterion evaluates the extent of automation, in respect to user involvement and interaction in the process. **9C.Introducing step-wise methodology for consolidation:** this assesses if a comprehensive method for systematic consolidation is provided. **10C.Supporting depth of research:** this general criterion identifies the level of research including conceptualization, formalization, implementation, and validation. The evaluation of each approach, as the answer to the RQ.1 question is summarized in Table 1. Please note that for the comparison purposes, the last row in this table only indicates a ranking of our own approach, latter addressed in Sect. 4) against our SLR criteria, as indicated with [**].

The assessment in Table 1 identifies the missing and challenging areas for further researches. For instance, columns 5C, 6C, 8C, and 9C represent aspects that need special attention in BP consolidation context. Consequently, introducing a collaboration-aware methodology and framework applicable to the VO context is valuable and lacking based on the analysis of results presented in columns 5C and 9C. Whereas, the evolution aspect of the consolidated model or regarding user involvement and preferences in the consolidation BP models through merge are other important subjects for research.

3.2 Stages in BP consolidation: results from SLR

Through the insights reached with our SLR study of the most relevant related research, we have identified the main aspects and steps in need of attention for development of a BP consolidation methodology. Five main stages are identified for this process, including: “*devise*”, “*formalize*”, “*merge*”, “*correct*”, and “*evolve*”. We discuss and analyze our findings related to each of these stages in this section. The five stages involved in a BP consolidation methodology development are as follows:

First stage is devising, which focuses on the planning of the aim and method of BP model consolidation process. The aim of BP consolidation is the first and foremost concern, since it specifies the intention behind the consolidation approach. For instance, two organizations in a merger may decide to standardize their BP models according to a specific reference model, while merging their BP models. Besides standardization, BP improvement [20,24] and configurable BP reference modeling [9,10,13] are some other well-known aims for consolidation, where the BP reference models represent best practice models that are formalized and address all aspects related to certain business domain [25].

One primary decision during the devise stage is related to the extent by which the BP models should be merged [3]. To facilitate making planning decisions for consolidation, two notions of: standardization and harmonization [24] are introduced in research. Standardization addresses merging of equivalent BP variants into one process model, while harmonization is applied to keep (not consolidate) a rationalized number of multiple equivalent BP models. These approaches in turn yield different methods of managing the BP variants [16]. Besides deciding on what plans to follow, devise stage also determines how to do the consolidation process by specifying the method of consolidation. The method mainly focuses on practical approaches for accomplishing the consolidation projects, e.g., being collaborative. In two principal scenarios of consolidation, either merger of two organizations, or the consolidation of two variant branches of a company, configuring of the organizational resources and the way they interact together is critical. This is addressed in research on organizational mergers [26].

Formalizing BP is the **second stage** of consolidation process, where the current BPs of an organization is represented by a BP modeling language, e.g., BPMN. The BP formalization has been researched extensively in BP modeling literature [5]. However, there are approaches in collaborative modeling [27] which facilitate this stage, but still there is no clear position for collaboration in BP consolidation process (See Table 1. Challenges related to formalizing BP models lie on: modeling at different levels of granularity, using different modeling languages, using synonyms and homonyms in labeling, and applying personal factors. In [28] BP merge formalization challenges are described and exemplified. While the above challenges are important, in many real-world projects, the main challenge is that formalized BP models do not exist [10,18].

Several studies have shown that large process models are more difficult to maintain and comprehend [25,29] while being error-prone, hence a large comprehensive model is not necessarily a favorite of modelers.

Third stage of BP consolidation process is **merging** the BP models. Our SLR shows that many consolidation approaches in BPM literature, address the (semi-)automated merging of the BP models. However, the merge stage includes three sub-stages. First, *similarity finding* is the primary step for merging BP models, i.e., finding equivalent BP models for combining them in the merge stage, which has received specific attention in literature. In [30] the similarity matching is classified into categories of label matching and BP models semantics matching. Due to different granularity in BP modeling, various labeling patterns, and heterogeneous terminology used in businesses, the similarity matching faces many difficulties and challenges [31].

The second sub-stage of BP model merge stage is focused on *resolving differences*. Following the identifica-

tion of matched/similar fragments, recognition of differences between the BP models is a main concern in BPM community. Thus, diagnosing differences, which is analyzed and categorized in [32], is the next step in BP consolidation roadmap. This step distinguishes the consolidation concept from the BP integration. In the integration of BPs, BP models collaboratively exchange information and messages; however, they do not change other partners' BPs. But in consolidation, planning the needed changes to address identified differences is a concern [19]. Without automation, it is tedious and time consuming for analysts to decide on differences and recognize common fragments in input BP models [10].

Combining BP models takes place at the third step of merging stage. Constituting the merged process model out of the input models is another research concern. Merging BP models is addressed in [9,10,13,14,20]. However, regarding the two perspectives of simple BP merge, and configurable BP merge, the mentioned approaches constitute the merged BP model generated from the input models of different merger organizations. When it comes to configurable BP merge, the (semi-)automated combining of BPs is very challenging because these approaches are syntactical and rarely represent the interest of business analysts if they are involved in redesigning BPs. Furthermore, the consolidation of all BPs in a configured merged model may result in a large and complex model [29], not favored by the users [25]. Moreover, lack of concrete evolving plan for such BP models may cause the merged process to divide it into multiple BP models [3], and provide more complexity in organizational communications, which again is not favored by users.

Forth stage of BP consolidation process is assuring the **correctness** of merged BP models [33]. During the automated merge of BP models, there is the risk of undesirable constructs to be generated; the so-called anomalies [5]. These anomalies may be in form of anti-patterns e.g., deadlock constructs [34], or ambiguous patterns e.g., OR-join constructs [35]. During the merge process, continuous validation checking is therefore necessary to avoid generation of ambiguous patterns. For instance in [36], one of the BP modeling guideline is to avoid generation of OR routing gateways, while in configurable BP merged models include typically OR gateways, since the OR gateway subsumes the behaviors of both the X-OR and the AND gateways. The correctness can also consider semantic aspects of BP modeling. For instance, due to the potential of contradictions in applied rules or behavior of the two input BP models from two organizations, there is a feasibility risk of constituting a correct merged BP model, e.g., one university gives priority to specific geographical zone in its admission procedure, while the other one is totally against any such discrimination.

Fifth stage of BP consolidation process is focused on managing the **evolution** of merged BP models. During the

BP evolution in an organization, the change patterns for achieving the desired evolved BP model are identified and propagated [37].

In principle, developing a solid methodology that can manage the above-mentioned stages of BP consolidation is a remarkable concern. A few methodologies are introduced that can manage BP model variants and even adapt them through their lifecycle [16]; however, there is still a lack of methodology for assuring the consistency of their consolidated BPs through merging their BP models.

3.3 Further SLR findings and analysis

Inputs to the BS consolidation process require being correct. The correctness is not just syntactical, but also behavioral, and otherwise the generated BSs will not properly execute. Therefore, a priority in developing BS consolidation methodology goes to preserving correctness, which is challenging, even due to different features of the BP modeling languages.

For instance, considering that BPMN as the de facto BP modeling language, it is often used for formalization of BPs in their consolidation process. But it is therefore necessary to also include an ambiguity resolution step to the devise stage of BS consolidation methodology. Furthermore, (semi-) automation is needed for it, since manual ambiguity resolution of BPMN models is itself not trivial.

Another issue to point out here is that even the merged models evolve over time. Consequently, the evolving stage of the consolidation methodology shall minimize the needed changes in the model while keeping the consistency of the merged model. These issues, as well as those addressed within the five above stages, constitute the main guidelines for design of our BP consolidation methodology.

4 A semiautomated BP merge methodology

This section addresses the concerns we formulated as RQ.2. We design a semiautomated methodology, guiding the process of consolidating BP models within a VO. The methodology consists of a set of practical steps that have to be performed using a tool chain in order to successfully introduce a collection of clean and valid consolidated processes for collaborating organizations. Please note that the steps in our methodology, while mostly automated, apply different methods varying between: manual, semiautomated and automated methods. All the mentioned relations are captured in Fig. 2 that illustrates the BP consolidation methodology flowchart. Further, each step is described in more detail in Sects. 4.1 to 4.6.

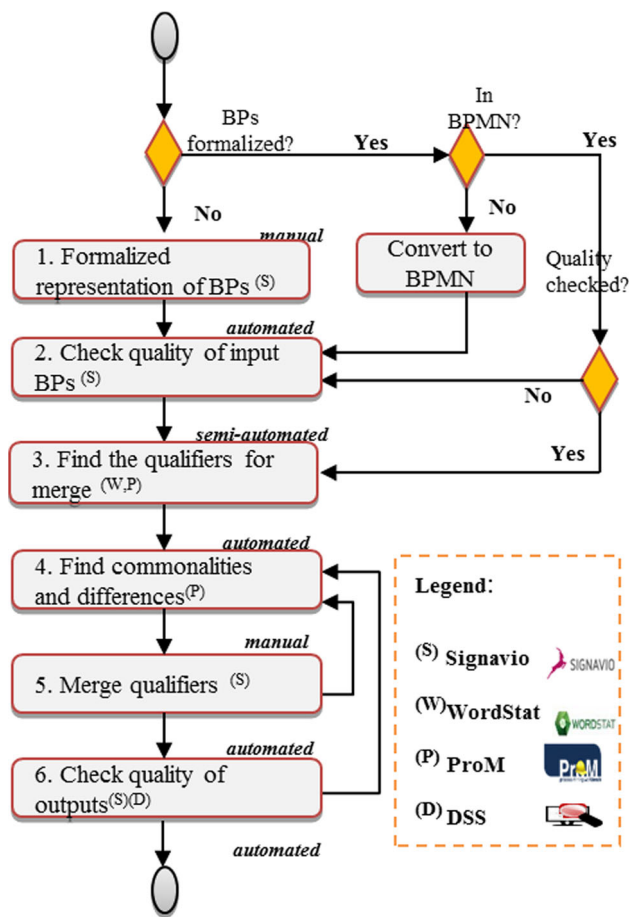


Fig. 2 BP consolidation methodology steps

4.1 Formalize representations of BPs

Each VO member organization performs a number of BPs within its individual operations. Those processes are not always formalized, and sometimes described only textually or in another ad hoc manner. For the purpose of processes' comparison, analysis, and consolidation, a collection of relevant BPs must be formalized in a consistent and common manner along all participating organizations. The formalization includes: process identification, process discovery, and process model design [38]. In real life, organizations formalize the design of their BPs in various ways, such as BPMN Diagrams or EPC process models. Different business process modeling languages (BPMLs) were evaluated in [39], who defined the Business Process Model and Notation (BPMN) [40] as the most suitable modeling language for use by VOs. Therefore, BPMN 2.0 is decided as the language used in our methodology. Variety of tools is available for designing the BPMN Diagram. We indicate the Web-based Signavio¹ software, as one of the top tools and most used to support the BP

¹ <http://www.signavio.com/>.

design, and to represent the executed processes. In summary, our formalization of BP representations in VOs includes the following sub-steps: *1. BP identification*—presenting a list of relevant BPs performed by all VO member organizations and providing the appropriate naming of the identified BPs, *2. BP discovery*—collecting the data regarding each BP and investigating it to properly describe the overview of actions performed in the specific BP, *3. BP design*—illustrating the described BP by designing its graphical model representation in BPMN.

4.2 Check quality of inputs

To apply the consolidation method, the input BP models must be syntactically correct. This syntactical correctness aspect indicates whether the representation of the BP model corresponds to the vocabulary and grammars of the BPMN 2.0 [40]. Moreover, the XML logs of the designed BP models must be verifiable by the software to assure the potential order of its execution. It means that the model is free of behavioral anomalies, such as deadlocks and/or livelocks [38]. Therefore, all the possible inconsistencies must be identified and removed at this early stage. Also for this step we use Signavio, as it incorporates the functionality that provides the check of BP model's correctness. Depending on a specific situation, the steps 4.1 and 4.2 might be omitted (e.g., if an organization has its BP models already formalized).

4.3 Find the qualifiers for merge

There may exist multiple similar BP models among various organizations. In case of forming a VO, we need to collect all its member organizations' BP models in a shared pool, which may contain redundancy occurrence and/or multiple BP variants. Firstly, we aim at identifying similarity of process in the VO pool and clustering similar BP models. Secondly, we enable further analysis for identification of outlier processes and cleaning the clusters. These sub-steps are further addressed below, after addressing two similarity checking approaches which we use.

To provide the more accurate semiautomated merge qualifiers identification, two kinds of preprocessing are employed. First kind is checking the *word similarity* on the names of processes. This approach produces a map drawn with all processes, where their closeness in distance from each other represents their similarity level. A function that displays the process names' map (illustrating the string semantic distances) is available in the WordStat software [41].

The *graph-based* similarity checking is the other kind of preprocessing. One of the algorithm that supports this idea and implemented in open-source software is introduced [31], provides the similarity calculation based on labeling, structural, and behavioral analysis of the input processes. This

algorithm is implemented in ProM 5.2² tool that is used in our methodology as the means to perform the graphical BP similarity check.

In the first sub-step, we apply the WordStat to the VO pool and produce a *word map* with all processes. We then manually draw circles around groups of processes identified as being closely related, through visual observation. In other words, here we are not searching only for the similar pairs of processes, but the potential groups of BP models that are similar with each other and can be merged into clusters. At the same time the wordmap indicates a number of BP outliers, which are the processes that are at far distance from those clustered.

However, by only evaluating the similarity among all gathered BP models, efficient identification of the merge qualifiers is not possible. For that purpose, more detailed sub-steps are followed:

- *Exclude BP outliers* This sub-step aims at retrieving a set of BPs, which stand highly individual in the wordmap, i.e., having below threshold results in the word similarity process. The individual processes will then go through a graph similarity check against the processes that are located within each word cluster. If still the similarity level are low, the resulted processes will be identified as the final set of potential outliers. The user will have the choice to decide on excluding them completely as an outlier. The main goal here is to potentially narrow down the number of BPs included in the analysis. The outliers will be considered for exclusion from any further similarity analysis.
- *Refine BP merge groups* This sub-step aims at taking each set of processes within one cluster, identified as a potential merge candidate groups, and analyzes them for formation of a final set of BPs to be merged. Here, the similarity is applied to identify if some processes do not properly belong to the group in order to finalize the merge candidate groups. This sub-step also requires validation by the user.

4.4 Find commonalities and differences

The particular BP models located within the final merge candidates at this step must then be compared with each other to enable the merge. Dijkman in [32] provides a technique to compare BP models and indicate their differences, which is implemented in ProM 5.2 software. However, such detailed analysis can only be performed only on pairs of two BPs. Therefore, we approach both the differences identification and merging activities in hierarchical steps. Based on discussion in [29], we start comparing pairs of the most similar

² <http://www.promtools.org/>.

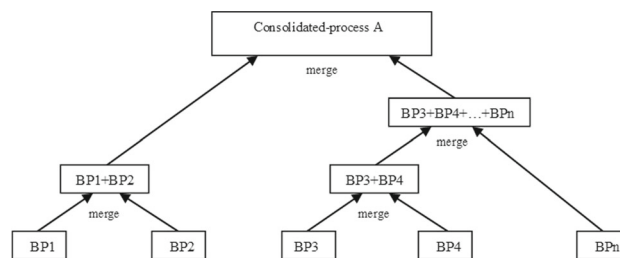


Fig. 3 Identifying differences in the merge hierarchy

BPs with each other, define their differences, and at the same time their commonalities (containing the same fragments), and merge the processes as described in the next step. Then, one pair of merged processes is next compared with another pair of merged processes (or a single BP), their commonalities and differences are examined, and they are merged at a higher hierarchical step. This iteration lasts until we reach the top level of the hierarchy and all the BPs in a merge candidate group are consolidation forming one final consolidated process. The abstraction of this approach is illustrated in Fig.3.

4.5 Merge qualifiers

Based on the information retrieved in the previous step, we can merge the processes by identifying the common fragments and resolving the differences. Some BP merging algorithms are available in research and implemented in an open-source software, e.g., “EPC merge” function introduced in ProM 5.2 by [9]. We evaluate the identified automated merging algorithm, by comparing the automatically obtained result with the desired manually created consolidated processes. Based on evaluation of this algorithm, after applying it on our construction case study, we decided to provide this task manually. The manual merging is performed mainly according to a method of La Rosa et al. [10] and is provided with the use of Signavio as the base environment. The main merging tasks include: (i) deciding on final labeling of equivalent nodes (e.g., deciding on using a more descriptive label), (ii) adding a node (preceded by “XOR” split) that is not present on one of the merged models, (iii) rearranging the structure of fragments, which represent different sequence; e.g., if two nodes are provided in two BP models in an opposite sequence, we rearrange them into a parallel structure preceded by “AND” split, (iv) adding gateways where necessary, e.g., if two BP models contain different nodes placed on the same position, both nodes should be included in a merged process, preceded by an added “XOR” gateway, (v) merging gateways, e.g., if two BP models contain different gateways (“AND” and “XOR”) placed on the same position, these gateways should be merged into an inclusive gateway (“OR”). The processes

are built upon each other, by adding the next BPs within the specific merge candidate group. After merging a pair of BP models, we return to step 4.4 to define the commonalities and differences of the next pair of BPs and merge this pair into one BP. This iteration lasts until all the BPs in a specific merge candidate cluster are merged. Consequently, we obtain a consolidated process including the behaviors of all its initial BPs.

4.6 Check quality of outputs

Similar to step 4.2, the syntactical quality of merged BP models are examined at this step, i.e., by Signavio automatic process quality check. The indicated errors and/or structure warnings must be reviewed, after which the process can be improved.

To obtain a clean process, the ambiguous BPMN patterns are also identified and reported. The ambiguous patterns are those workflow patterns [42] which includes OR-join and complex-join (e.g., General Synchronizing Merge, Blocking Partial Join, and Structured Discriminator) and face problems during execution with BP management systems. A discussion on ambiguity and inconsistency with workflow engines could be found in [5,35,42]. To identify ambiguous workflow patterns, we apply our own developed tool, named Disambiguation Decision System (DSS). The tool is a subject of another research.

After this step, we can proceed to consolidate the next merge candidate group of BP models, by returning to step 4.4. This iteration lasts until we obtain a complete set of consolidated processes.

5 Evaluation of BP consolidation methodology

In this section RQ.3 is addressed by applying and thus evaluating the methodology introduced in Sect. 4 on a real large example case from building industry. This exercise proved fruitful, and the lessons learned and encountered limitations during this practical exercise provide the base for future research in this area. This industry case concerns 10 SMEs operating in the Polish market, as indicated in lowest rectangle in Fig. 4. The figure provides a context model for their typical interactions in a construction project.

A commonly used name for customer in building industry is “investor”. The investor is usually not a specialist in a construction domain, so the help of control inspectors is often needed during the project. Control inspector is a source of practical information for the architect and the investor and is a representative of the investor on the construction site, controlling the quality of provided services. At present, contractors operate individually at the construction site. While

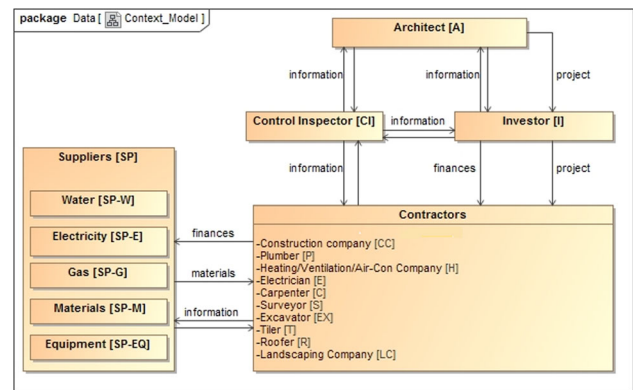


Fig. 4 Construction case study context model

the same group of companies usually work together in similar projects, there is currently no long-term agreement for collaboration among the involved organizations. Therefore, the situation does not support process sharing and delivering value-added services by them. However, the construction industry remains increasingly competitive, and difficult to survive in, especially for SMEs that are analyzed in our case study. They are continuously surpassed by large construction companies that offer fast services, which are cheaper but with lower quality. Therefore, the emergence of niche high-quality construction projects enables SMEs collaboration through VOs, to increase their profitability, survival rate, and seizing more market opportunities.

To examine the applicability of the methodology, we design our case study, where the construction company plays the pivotal roles and coordinates the interactions with other companies. The required data are collected from the existing documents or during the interviews with representatives of companies. To develop BPMN models a moderator was involved. Moreover, the results at different stages of the case study were evaluated and accepted by Construction Company and other involved organizations. In the remaining of this section we briefly address in Sects. 5.1 to 5.6, the application of the steps introduced in Sect. 4 for our BP consolidation methodology, to the real case study of ten collaborating contractor companies, as represented in Fig. 4. The complementary information of case study is available on “<http://files.fcngroup.nl/BPCON/>”.

5.1 Formalize representations of BPs

For the ten contractor members in this VO, a total number of 82 BPs are identified. These BPs are either extracted from the existing documents of companies’ BPs (e.g., textual descriptions or flowcharts), or identified through interviews with the representatives of companies. One example such process, for **supplying materials** performed by the electrician company [E], is presented in Fig. 5 and preceded by its textual descrip-

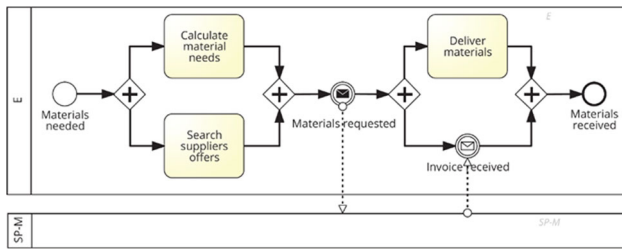


Fig. 5 E2 process formalization(“SP-M- supplier of materials”)

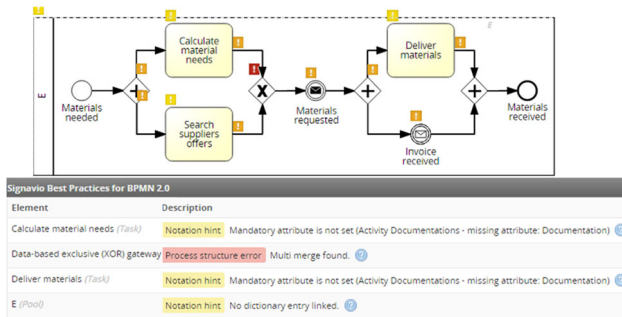


Fig. 6 Quality check by Signavio—process with error and warnings next to XOR

tion as follows. *Electrician—supply materials [process E2]:* “Supplying starts when a need for materials appears. Then, electrician calculates the amount of needed materials, in parallel with searching other suppliers offers. As an outcome of these tasks, a materials request is created and sent to another specific supplier. In the end, both the requested materials and an invoice are delivered and the process ends”. The results were further confirmed by the representative of the companies.

5.2 Check quality of inputs

All designed BPs are quality checked by Signavio software for potential syntactic problems. As an example, imagine that by mistake for E2 we generate the BP in Fig. 6. Running Signavio on it indicates some problems found for Electricians supplying materials process (E2), as indicated in Fig. 6, with “!” next to the XOR operation (in the center). The cleaned process is then produced to correct syntactical errors, as in Fig. 5.

5.3 Find the qualifiers for merge

Analyzing the formalized 82 BP models can spot several BPs that are more or less repeated by a number of organizations within this VO. Identification of the so-called merge qualifier processes is the necessary first step in BP consolidation, as addressed in Sect. 4.3. Note that only cross-organizational BP similarity is checked and not intra-organizational similarities.

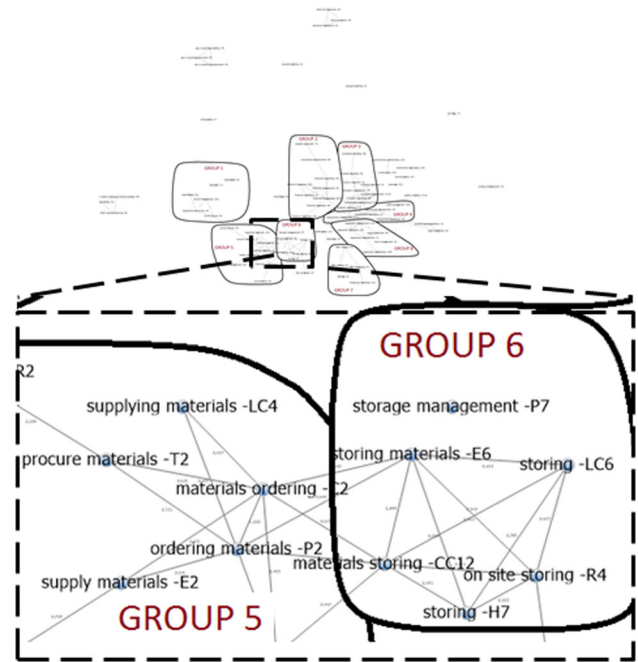


Fig. 7 Process names word map and clustering

5.3.1 Exclude BP outlier

Using the Link Analysis function (in WordStat) a word map is created visualizing all similarities. Using the map, we first manually cluster potential candidate BPs for merge, as presented in Fig. 7. Please note that while illustrated processes are comprehensive, a part of Fig. 7 is enlarged to be readable. Each of the eight groups indicates a potential candidate for clustering and the rest are 27 potential outliers among the 82 BPs. Running the automated similarity checking, based on [31] identify that 27 out of 82 processes are different from 93.3 % of the cross-organizational BP models (Considering similarity threshold of 35%). Hence, produced results indicate BP outliers and are excluded from further steps of this methodology.

5.3.2 Refine BP merge group

Eight groups are manually created on the results from running of the two automated similarity checks addressed above. In order to validate these clusters for merge purposes, we run automated graph similarity calculations once within each group, and once outside of these groups, with the use of ProM 5.2 tool. This task however requires a double conversion of the BPMN models (through Petri Net into EPC), since the automated comparison tool is only available for the EPC models, and to the best of our knowledge, there are no open-source software that offers a single conversion between the BPMN and EPC languages. In practice, the first conversion (into Petri Net) was provided in ProM 6.4, and the second

Table 2 Merge candidate groups' results overview

Merge candidate group	Similar inside group%	Dissimilar inside group%	Similar outside group%	Dissimilar outside group%
1	57.14	42.86	26.43	73.57
2	100	0	12.50	87.50
3	77.14	22.86	13.75	86.25
4	100	0	23.56	76.44
5	61.43	38.57	8.00	92.00
6	100	0	22.47	77.53
7	55.00	45.00	10.76	89.24
8	86.67	13.33	19.00	81.00

conversion (into EPC) was done in ProM 5.2. Provided in ProM, Petri Net as a formal language accurately represents the causal dependencies of BPMN and EPC process models [43]. We are certainly aware of the fact that the double conversion may cause some data loss and thus may influence the results. Nevertheless, for the similarity evaluation, we used a threshold of 35%, even higher than the acceptable estimation stated in [44] that uses 33% similarity threshold. Table 2 shows positive (above threshold) and negative (below threshold) similarity results, inside and outside each merge candidate groups. Based on results of Table 2, we assess the usability of the proposed method. We can notice that inside all groups, the rates of their BP models being similar to any of the BP models outside of merge candidate groups (indicated in the second column from right) is significantly lower than the rates of BPs that are not similar (indicated in the first column from right). In fact, the positive similarity with BPs outside the group is measured mostly under 25%. Therefore, the method proves its applicability, by showing lack of similarity with other BPs, at a satisfactory level. By analyzing the left part of Table 2, we can evaluate the internal group BPs and their candidacy for integration within the specific group. Furthermore, checking similarities inside each group, the similarity among BPs is larger than 50% and sometimes 100% (indicated in the fourth column from right). The main aim here is to identify and suggest to users groups of BPs as candidates for merging and those that are outliers in this process. To assess the internal structure of each group, we must look closer into individual BP graph similarity check results. Groups 2, 4, and 6 show completely similarity internally, therefore they represent very good candidate for BP clustering. So we then analyze only each remaining group manually by domain experts, examining details of similarity checks among all their BPs, and deciding if they can or cannot be merged.

5.4 Find commonalities and differences

With automated software, we identify differences in the merge hierarchy for each of the eight groups, and resolve

Table 3 Group 5 similarity results

	R6	R2	T2	C2	E2	H2	P2
CC3	0.45	0.21	0.42	0.44	0.49	0.35	0.44
P2	0.33	0.37	0.65	0.75	0.61	0.41	
H2	0.33	0.25	0.56	0.43	0.37		
E2	0.41	0.27	0.65	0.77			
C2	0.42	0.23	0.80				
T2	0.33	0.37					

difference with manual help from experts. To give an example, we illustrate the similarity table of group 5 in Table 3, which is calculated in the previous step (See 5.3). Consider merging of two BPs in group 5: C2 for “materials ordering”, and T2 for “procure materials”, which are identified as being highly similar (identified with 80% similarity, in previous step), only 4 differences are identified in them by the software.

Analysis performed in ProM software indicates only four differences between these BPs (“Difference [number]” corresponds to a specific type of difference, as classified in [32] as follows: (i) *Difference 31*: Skipped activity; transition ‘_join’ from provided behavior has no equivalent in required behavior (ii) *Difference 11*: additional dependencies; transition ‘Materials requested’ from provided behavior has additional dependencies with respect to transition ‘Materials requested’ from required behavior on ‘Search suppliers offers+complete’ (iii) *Difference 31*: Skipped activity; transition ‘_fork’ from provided behavior has no equivalence in required behavior (iv) *Difference 21*: Different moments; transition ‘Calculate material needs+start’ from provided behavior [in T2] occurs at a different moment than transition ‘Calculate material needs+start’ from required behavior [in C2].

The above-mentioned list points out some structural differences, e.g., not matching nodes sequence and additional fork and join existing in only one of the presented BPs. These differences are very simple, because the analyzed BPs are

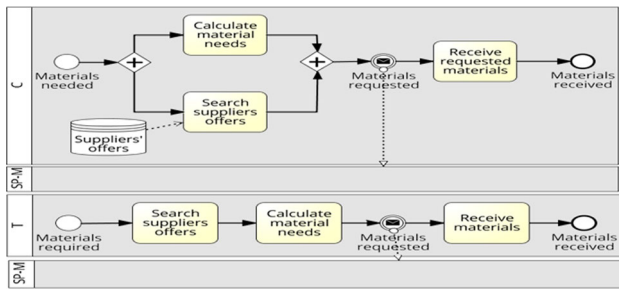


Fig. 8 BP models C2 (above) and T2 before merging

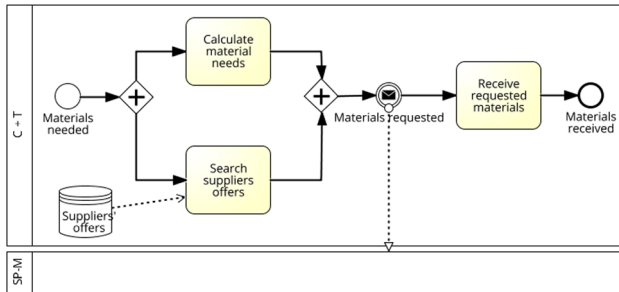


Fig. 9 Merged model C2 + T2

highly similar, as can be noticed by checking them in Fig. 8. For comparison of less similar BPs, the list of differences will grow much longer. For example, comparing E2 and P2, for which similarity value is 0,61, there are 11 differences identified, whereas for C2 and T2 with similarity of 0,80, only 4 differences are identified. However, in some cases a big part of differences concerns the structural part of one BP model, which is not present in the other BP, and this can be easily adjusted by adding a configurable gateway and inserting the entire fragment, which is missing in the second BP.

5.5 Merge candidate

At this stage, common fragments are identified and differences are resolved manually to reach the merged BP, as described in Sect. 4.5. The manual merge is, however, provided step-wise, according to the merging hierarchy developed in Sect. 5.4. For the introduced example case (merging C2 and T2), we can notice that the process C2 actually contains the structure of process T2. The activities of “Searching suppliers offers” and “Calculating material needs” that are organized in a sequence in T2 can be therefore added to that sequence using additional AND split and join for them. However, for the needs of process C2, this sequence can be executed in parallel, as illustrated on its initial model. Concluding this merging instance, the process C2 + T2 generates exactly the same as process C2 (See Fig. 9).

The same procedure applies to other processes in the cluster, as also shown in Table 3. This means, in the next round, merging E2 and P2. Furthermore, the two merged processes of “E2+P2” and “C2+T2” are merged. Similarly, the two BP models of “CC3” and “R6” are merged with the previous merged models. In a similar manner, the two merged models of “H2+R2” and “C2+T2+E2+P2+CC3+R6” are merged at the last stage, and make the final consolidated process of group 5, illustrated on Fig. 10. Note that, as mentioned in Sect. 4.5 we aimed at including all the nodes, structures and behaviors of the initial BPs. However, some additional decisions were made by expert decision makers, e.g., regarding removal of a node that should not have been included in the model.

5.6 Check quality of outputs

Since our consolidated processes are also created in Signavio software, we provide the quality (i.e., the syntactical correct-

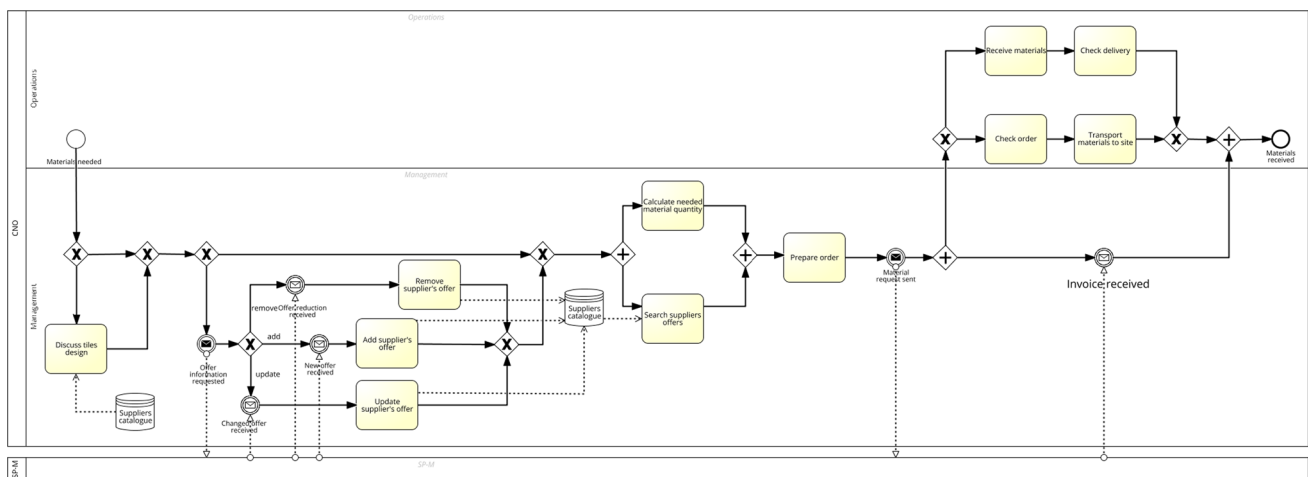


Fig. 10 Consolidated process 5

ness) check the same way as introduced for the Sect. 4.6. In general, the consolidated processes are complex, so there is a higher risk of generating certain gateway inconsistencies, e.g., causing deadlock errors in the resulted integrated BP. Therefore, it is essential to check whether any resulted process includes such errors or ambiguities and if so, to disambiguate them. Semiautomated disambiguation of BP models is also conducted. Since this part is the subject of one of our ongoing research area, and will be reported in forthcoming publications. After disambiguation step the following merged model is finalized by domain experts.

5.7 Lessons learned

The main lessons we learned through conducting the case study are summarized, as follows:

- There is no open-source tool that efficiently provides the functionalities (i.e., finding commonalities and differences) needed to merge BPMN Diagrams. Nevertheless, to work with the tools introduced in our methodology, familiarity with BPMN would be sufficient for consolidating BP models.
- Conducting the “merge qualifiers” step, defined in Sect. 4.5 was not a trivial job. The coordination of representatives from each involved organization needed special attentions. Several issues caused that, including the different levels of modeling knowledge backgrounds and heterogeneities of the initial model representations (e.g., textual descriptions, flowcharts, and ad hoc notations). In this regard, a moderator assigned to assist in this step.
- To find better similar BP names automatically, providing a proper naming that is understandable by the software is necessary. We found that using any acronyms and special characters in BP names influences the outcome of word map (produced by WordStat). Therefore, we provided the complete and clear names of BPs. For instance, in our case study; we had to change the BP names and acronyms including the “mgmt.” into full names, i.e., “financial management”. Meanwhile, a glossary that included the definitions of the common business terminologies was also provided to help in the consolidation process.
- We obtained the similarity by comparing complete BP models, and not BP fragments. Some organizations designed their BP models in different levels of granularities. For example, in one company choosing a supplier in a BP model showed by several steps, including “Check vendor’s black list” and “Pre-check supplier’s inventory”, while in other company they represented this function in a single task. This issue, in turn, affected the outputs of the tools that mostly rely on word similarity checking. Nevertheless, we decided on the appropriate level

of granularity in the consolidated BPs by regarding the suggestions received from the VO members.

- The OR gateway in BPMN can subsume both mutually exclusive and parallel behaviors. In other words, this gateway allows choosing one or all possible alternatives paths in a process. Accordingly, to combine multiple functions in a consolidated process, this gateway was used by VO representatives. For instance, the construction company could use either its internal store, or rent an external store, or could use both internal and external stores in a project. Although OR gateway was quite practical in developing consolidated BPs, some OR included fragments (e.g., OR-join) in the consolidated BP models faced difficulties in execution, due to their ambiguous semantics [35,42]. Manually identifying and removing the problematic OR-join fragments was also a difficult job for VO representatives.

6 Conclusion and discussion

In relation to RQ.1, we performed an extensive SLR on the topic of BS/BP consolidation through merge. The results show gaps in a methodology that: (i) addresses the formalized requirement of VOs and values user preference in developing the merged BS/BP model, (ii) reduces the time and efforts of manual merging of BP models, and serves as the base for automating the consolidation procedure in the VO context. The five main stages of: devise, formalize, merge, correct, and evolve are identified through SLR as being needed for development of a methodology for BS/BP consolidation through merging. The first four stages are then addressed in the paper during the design of our methodology, which is the subject of RQ.2 (addressed below). However, the evolving stage is planned to be developed as future work, and is not examined with practically.

The RQ.2 is addressed in the paper through designing a merged-based BS consolidation methodology, which guarantees the successful co-creation of a clean (i.e., anomaly-free consolidated model) and valid (i.e., approved by domain experts) consolidated process among member organizations in the VO. The steps needed to create the set of consolidated processes, as indicated in the methodology, include: formalizing representations of BPs, checking the quality of inputs, finding the qualifiers for merge, finding commonalities and differences of the merge qualifiers, merging the groups of BPs, and checking quality of outputs. Note that the ranking of our proposed approach against the SLR criteria is also provided in the last row of Table 1, indicated with [**]. This work not only addresses the lack of a methodology to support BP consolidation needs in VO context but also considers the executability of the consolidated BP models.

In relation to RQ.3, the created methodology was applied to a real case study of a VO for constructing eco-friendly buildings. This extensive case evaluated the applicability and validity of our introduced steps for the methodology by the involved organizations. The VO members were satisfied by the consolidated BP models and used them. Although our findings are yielded from a single-case study, they are useful for other VOs that intend to apply the suggested methodology. The methodology fits better with the VOs that aim at establishing a long-term collaboration. Furthermore, having more formalized BP models in BPMN saves time in the consolidation process. In the case study, we defined an important role of moderator that assists in BPMN modeling; nevertheless, familiarity of VO representatives with modeling knowledge could also be an advantage to the successful use of the introduced methodology.

As a part of our future work, we address the automation of a currently manual step in our BS consolidation methodology. The fifth step of the methodology is currently a manual process of BP merging, where process supported by a BP modeling editor. Accordingly, it is much more desirable if this step can also be (semi-)automated. In other words, it is desirable to have an automated good quality tool that can be uniformly applied to all BPs that are highly similar within the VO's shared pool of business services.

Further, more complex disambiguation of the merged BSs and checking for the executability of the merged results, e.g., by identifying deadlocks, livelocks or handling the ambiguous workflow patterns (e.g., OR-join) in the produced results of the automated merge, are the next challenging steps of the research presented in this paper which requires more research on the subject.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Afsarmanesh H, Camarinha-Matos LM, Msanjila SS (2009) On management of 2nd generation virtual organizations breeding environments. *Ann Rev Control* 33(2):209–219
- Camarinha-Matos LM, Afsarmanesh H, Galeano N, Molina A (2009) Collaborative networked organizations-concepts and practice in manufacturing enterprises. *Comput Ind Eng* 57(1):46–60
- Dumas M (2011) Consolidated management of business process variants. In: *International conference on business process management*. Springer, pp 1–1
- Camarinha-Matos LM, Afsarmanesh H, Oliveira AI, Ferrada F (2013) Cloud-based collaborative business services provision. In: *International conference on enterprise information systems*. Springer, pp 366–384
- van der Aalst WM (2013) *Business process management: a comprehensive survey*. ISRN software engineering 2013
- Jones S, Mandyck J (2016) *World green building trends smartmarket report*. Technical report, prepared for McGraw-Hill Construction and United Technologies Climate, Controls & Security
- Cheng JC, Law KH, Bjornsson H, Jones A, Sriram R (2010) A service oriented framework for construction supply chain integration. *Autom Constr* 19(2):245–260
- Kitchenham B, Brereton P (2013) A systematic review of systematic review process research in software engineering. *Inf Softw Technol* 55(12):2049–2075
- Gottschalk F, Wagemakers TA, Jansen-Vullers MH, van der Aalst WM, La Rosa M (2009) Configurable process models: experiences from a municipality case study. In: *International conference on advanced information systems engineering*. Springer, pp 486–500
- La Rosa M, Dumas M, Uba R, Dijkman R (2013) Business process model merging: an approach to business process consolidation. *ACM Trans Softw Eng Methodol (TOSEM)* 22(2):11
- Yin RK (2013) *Case study research: design and methods*, 5th edn. Sage publications, Thousand Oaks
- Hevner A, Chatterjee S (2010) *Design science research in information systems*. Springer, Berlin
- Küster JM, Gerth C, Förster A, Engels G (2008) A tool for process merging in business-driven development. In: *CAiSE forum, Citeseer*, vol 344
- Gerth C (2013) *Business process models*. Change management, Lecture Notes in Computer Science, vol 7849. Springer
- Schunselaar DM, Verbeek E, van der Aalst WM, Raijers HA (2012) Creating sound and reversible configurable process models using cosenets. In: *International conference on business information systems*. Springer, pp 24–35
- Ayora C, Torres V, Weber B, Reichert M, Pelechano V (2015) Vivace: a framework for the systematic evaluation of variability support in process-aware information systems. *Inf Softw Technol* 57:248–276
- La Rosa M, van der Aalst WM, Dumas M, Milani FP (2013) Business process variability modeling: a survey. *ACM Comput Surv* 50:2
- Dijkman RM, La Rosa M, Reijers HA (2012) Managing large collections of business process models-current techniques and challenges. *Comput Ind* 63(2):91–97
- Sun S, Kumar A, Yen J (2006) Merging workflows: a new perspective on connecting business processes. *Decis Supp Syst* 42(2):844–858
- Mendling J, Simon C (2006) Business process design by view integration. In: *International conference on business process management*. Springer, pp 55–64
- Reijers HA, Mans R, van der Toorn RA (2009) Improved model management with aggregated business process models. *Data Knowl Eng* 68(2):221–243
- Li C, Reichert M, Wombacher A (2010) The MinAdept clustering approach for discovering reference process models out of process variants. *Int J Cooperative Inf Syst* 19(03n04):159–203
- Derguech W, Bhiri S (2011) Epc. In: *International conference on business information systems*. Springer, pp 86–97
- Romero H, Dijkman R, Grefen P, van Weele A (2011) Harmonization of business process models. In: *International conference on business process management*. Springer, pp 13–24
- Recker J, Rosemann M, Van Der Aalst W (2005) On the user perception of configurable reference process models-initial insights. *ACIS 2005 proceedings* p 66
- Gomes E, Angwin DN, Weber Y, Yedidia Tarba S (2013) Critical success factors through the mergers and acquisitions process: revealing pre-and post-m&a connections for improved performance. *Thunderbird Int Bus Rev* 55(1):13–35

27. Rittgen P (2010) Success factors of e-collaboration in business process modeling. In: International conference on advanced information systems engineering. Springer, pp 24–37
28. Malekan HS, Mehrizi MHR, Afsarmanesh H (2016) Positioning collaboration in business process model consolidation in vos. In: Working conference on virtual enterprises. Springer, pp 110–123
29. Vogelaar J, Verbeek H, Luka B, van der Aalst WM (2011) Comparing business processes to determine the feasibility of configurable models: a case study. In: International conference on business process management. Springer, pp 50–61
30. Becker M, Laue R (2012) A comparative survey of business process similarity measures. *Comput Ind* 63(2):148–167
31. Dijkman R, Dumas M, Van Dongen B, Käärrik R, Mendling J (2011) Similarity of business process models: metrics and evaluation. *Inf Syst* 36(2):498–516
32. Dijkman R (2007) A classification of differences between similar BusinessProcesses. In: 11th IEEE International enterprise distributed object computing conference, 2007. EDOC 2007. IEEE, pp 37–37
33. van der Aalst WM, Dumas M, Gottschalk F, Ter Hofstede AH, La Rosa M, Mendling J (2010) Preserving correctness during business process model configuration. *Form Asp Comput* 22(3–4):459–482
34. Koehler J, Vanhatalo J (2007) Process anti-patterns: how to avoid the common traps of business process modeling. *IBM WebSphere Dev Tech J* 10(2):4
35. Börger E (2012) Approaches to modeling business processes: a critical analysis of BPMN, workflow patterns and YAWL. *Softw Syst Model* 11(3):305–318
36. Mendling J, Reijers HA, van der Aalst WM (2010) Seven process modeling guidelines (7pmg). *Inf Softw Technol* 52(2):127–136
37. Weber B, Reichert M, Rinderle-Ma S (2008) Change patterns and change support features-enhancing flexibility in process-aware information systems. *Data Knowl Eng* 66(3):438–466
38. Dumas M, La Rosa M, Mendling J, Reijers HA (2013) Fundamentals of business process management, vol 1. Springer, Berlin
39. Malekan HS, Afsarmanesh H (2013) Overview of business process modeling languages supporting enterprise collaboration. In: International symposium on business modeling and software design. Springer, pp 24–45
40. OMG (2013) Notation (bpmn) version 2.0.2. OMG specification, Object Management Group
41. Pollach I (2010) Software review: Wordstat 5.0. Organizational research methods
42. Russell N, van der Aalst WM, ter Hofstede AH (2016) Workflow patterns: the definitive guide. MIT Press, Boston
43. van der Aalst WM, van Dongen BF, Günther CW, Mans R, de Medeiros AKA, Rozinat A, Rubin VA, Song M, Verbeek HE, Weijters A (2007) Prom 4.0: comprehensive support for real process analysis. *ICATPN* 4546:484–494
44. Sirageldin A, Selamat A, Ibrahim R (2011) Graph-based simulated annealing and support vector machine in malware detection. In: 2011 5th Malaysian conference in software engineering (MySEC). IEEE, pp 361–364