MDPI

Article

Eligibility of BPMN Models for Business Process Redesign

George Tsakalidis [1] **, Kostas Vergidis** [1,*] **, Georgia Kougka** [2] **and Anastasios Gounaris** [2]

1 Department of Applied Informatics, University of Macedonia, Thessaloniki 54636, Greece
2 Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece
* Correspondence: kvergidis@uom.gr; Tel.: +30-2310-891-637

check for updates

Abstract: Business process redesign (BPR) is an organizational initiative for achieving competitive multi-faceted advantages regarding business processes, in terms of cycle time, quality, cost, customer satisfaction and other critical performance metrics. In spite of the fact that BPR tools and methodologies are increasingly being adopted, process innovation efforts have proven ineffective in delivering the expected outcome. This paper investigates the eligibility of BPMN process models towards the application of redesign methods inspired by data-flow communities. In previous work, the transformation of a business process model to a directed acyclic graph (DAG) has yielded notable optimization results for determining average performance of process executions consisting of ad-hoc processes. Still, the utilization encountered drawbacks due to a lack of input specification, complexity assessment and normalization of the BPMN model and application to more generic business process cases. This paper presents an assessment mechanism that measures the eligibility of a BPMN model and its capability to be effectively transformed to a DAG and be further subjected to data-centric workflow optimization methods. The proposed mechanism evaluates the model type, complexity metrics, normalization and optimization capability of candidate process models, while at the same time allowing users to set their desired complexity thresholds. An indicative example is used to demonstrate the assessment phases and to illustrate the usability of the proposed mechanism towards the advancement and facilitation of the optimization phase. Finally, the authors review BPMN models from both an SOA-based business process design (BPD) repository and relevant literature and assess their eligibility.

Keywords: business process redesign; complexity assessment; normalization; data-centric workflow optimization

1. Introduction

A business process (BP) is considered a collection of interrelated activities, events and decision points involving a number of actors and objects, orderly performed to actualize an outcome of value to at least one customer [1]. In order to respond to increasingly volatile markets, companies are examining the dynamic redesign of their core BPs to improve performance metrics and adaptability. Through business process redesign (BPR), organizations implement innovative changes to establish improvements to critical success factors [2].

In [1], Dumas et al. classify redesign techniques into two categories: (a) heuristic process redesign, which builds upon an extensive set of redesign options to refine an existing process; and (b) product-based design, which refers to the radical re-engineering of processes and (in essence) replacing rather than refining the existing ones. Some examples are: customer heuristics, which consider giving control to the customer, reducing the number of contacts between the organization and the customer, and allowing the customer to be more involved in the whole process; BP operation

heuristics, which may eliminate redundant activities, decide between case-based and batch work modes, provide alternative task implementations, and compose smaller activities rather than larger ones; and BP behavior heuristics, which regulate the logic within the BP through applying the practices of resequencing, parallelism, knockout, and exception. Another redesign classification scheme is presented in [3], through a quantitative manner that exploits queue theory to emulate the behavior of facilities (either human or machine) that execute tasks. The problems covered include: (a) the exact task ordering (e.g., whether a strict or partial task execution ordering is chosen); (b) resource allocation (e.g., whether subsequent tasks can be allocated to the same facility or not); (c) whether a specific facility takes management decisions at runtime; and (d) task implementation and considering the possibility of processing a specific task in a specialized rather than a generic manner. Finally, La Rosa et al. [4] introduced the notion of re-configurable processes where, at some point, different tokens in the same application may follow different paths. This notion encompasses modifications in ordering, task implementation and resource allocation, but no exact holistic technique is proposed.

In spite of the fact that a thorough analysis of a BP typically sparks various ideas and directions for redesign, it is often approached in a non-systematic way, and is largely considered to be a creative activity. In this sense, parts of the wide-ranging spectrum of potential redesign options could be omitted, as in the majority of creative techniques [1]. The latter enhances the need for narrowing the redesign perspective and specifying the particular aspects and capability of a method to yield better redesign options. For instance, a BP can be reformed to optimize metrics like resource utilization, maximal throughput and flow time, and thus heuristics like task resequencing, parallelism, and task composition have to be applied [5]. Given many such alternatives, there is no systematic way to evaluate these heuristics and the tradeoffs among them. In [6], Kumar and Indradat describe a novel approach based on a formal model for optimizing the benefits or net effects of a redesign practice regarding a baseline design, and illustrate how it can be used to evaluate and compare alternative models in terms of both design and runtime. A limitation is that the user must identify all valid redesign strategies and manually include them in the model, which entails that more semantic information about the process model may be needed by the system. In [7], a flexible framework is proposed for modeling, simulating and reengineering BPs in a cost-effective way. Due to the fact that, in most cases, there are more than one redesign option, these are evaluated against the anticipated benefits and strategic objectives of the organization. The optimal ones are manually selected and further analyzed to identify neglected problems in a procedure that comprises a BPR methodology. A novel approach presented in [8] aims to provide automated performance optimization of the BP execution by transferring knowledge from the related field of data-centric workflows. This is achieved through a transformation process of an initial BPMN model to an equivalent annotated directed acyclic graph (DAG) and the optimization of the average performance of process executions. Data-flow optimization can lead to performance improvements, and more importantly, has the potential to lift the burden of manually fixing all implementation details from the data-flow designers (a key motivation for automated optimization solutions [9]).

An important factor that is overlooked in the majority of such approaches is the lack of an evaluation mechanism for the BP models to be redesigned. In the majority of existing redesign approaches, there is no consideration of the model type, its complexity level, the overall redesign feasibility, or, most importantly, the applicability of particular redesign heuristic(s). One of the obstacles for evaluating BPs is their complexity [10], and new methodologies are needed to handle this complexity, especially in terms of how to integrate process information into enterprise networks [11]. In this paper, the authors introduce an evaluation mechanism for BPMN models that is motivated by this research gap and the adversities encountered in previous work [8]. This is achieved through a set of criteria that emulates a decision support system and assesses essential model characteristics like complexity, redesign capability and normalization, to render a model eligible for further transformation and optimization. The proposed mechanism can also be potentially used for supporting the creation of reusable component BP models. In [12], a set of activities, including the execution of a decomposition

algorithm, creates a component model repository that facilitates process redesign and improvement. The method uses complete BPMN diagrams as input, and the repository of reusable parts follows BP modeling guidelines to avoid typical anomalies, which is a common objective with the assessment mechanism put forward in this paper.

This paper is structured as follows: Section 2 puts forward the concept of eligibility assessment for BPMN models by proposing a set of criteria, and showcases through a BPMN example how complexity assessment, normalization and transformation steps are designed towards optimization. Section 3 reviews the eligibility of 20 BPMN input models extracted from the SOA-based business process database and relevant literature. Section 4 discusses the paper findings and ideas for future work.

2. Eligibility Assessment of BPMN Models

The diversity of BPMN constructs and semantics can produce a range of models varying in complexity and purpose that can be challenging to standardize for a formal optimization framework. Motivated by the research issues mentioned in the previous section, this paper proposes a set of eligibility criteria that formalize a BPMN model and facilitate the transformation and optimization process by providing reliable and consistent input. The proposed assessment mechanism focuses on a particular set of criteria:

- the input model type;
- the features of the model that allow optimization (i.e., resequencing capability);
- the structuredness of the model;
- the model complexity.

Based on the above criteria, the authors consider an eligible BPMN input model to be one that models a structured, private, non-executable BP with a resequencing capability. The "Boarding Procedure model" (Figure 1) from the SOA-based business process database [13] is selected as an indicative example to illustrate the assessment mechanism in detail. The main phases of the eligibility assessment (Figure 2) follow.

2.1. Model Type Check

BPMN is designed to cover many types of modeling and allows the creation of end-to-end BPs. There are three model types within an end-to-end BPMN model, namely Processes (Orchestration), Choreographies and Collaborations [14]. The three basic types of Processes (Orchestration) are: (a) private non-executable (internal) BPs; (b) private executable (internal) BPs; and (c) public processes. A public process represents the interactions between a private BP and another process or participant. Only those activities that are used to communicate with other participant(s), plus the order of these activities, are included in the public process. All other "internal" activities of the private BP are not shown in the public process. Private BPs, also known as workflows or BPM processes, are internal to a specific organization and are divided into executable and non-executable processes. At this step, the input BPMN model is checked for whether it corresponds to a private non-executable BP. These models document process behavior at a modeler-defined level of detail, hence, information needed for execution, such as formal condition expressions, are typically not included. The example BPMN model (Figure 1) is a private BP that illustrates a process internal to an airline company. Due to the fact that information needed for execution are not available, such as formal condition expressions, the BPMN model depicts a private non-executable BP. In future work, the authors intend to use private executable BPs aiming for BP automation through automated performance optimization.
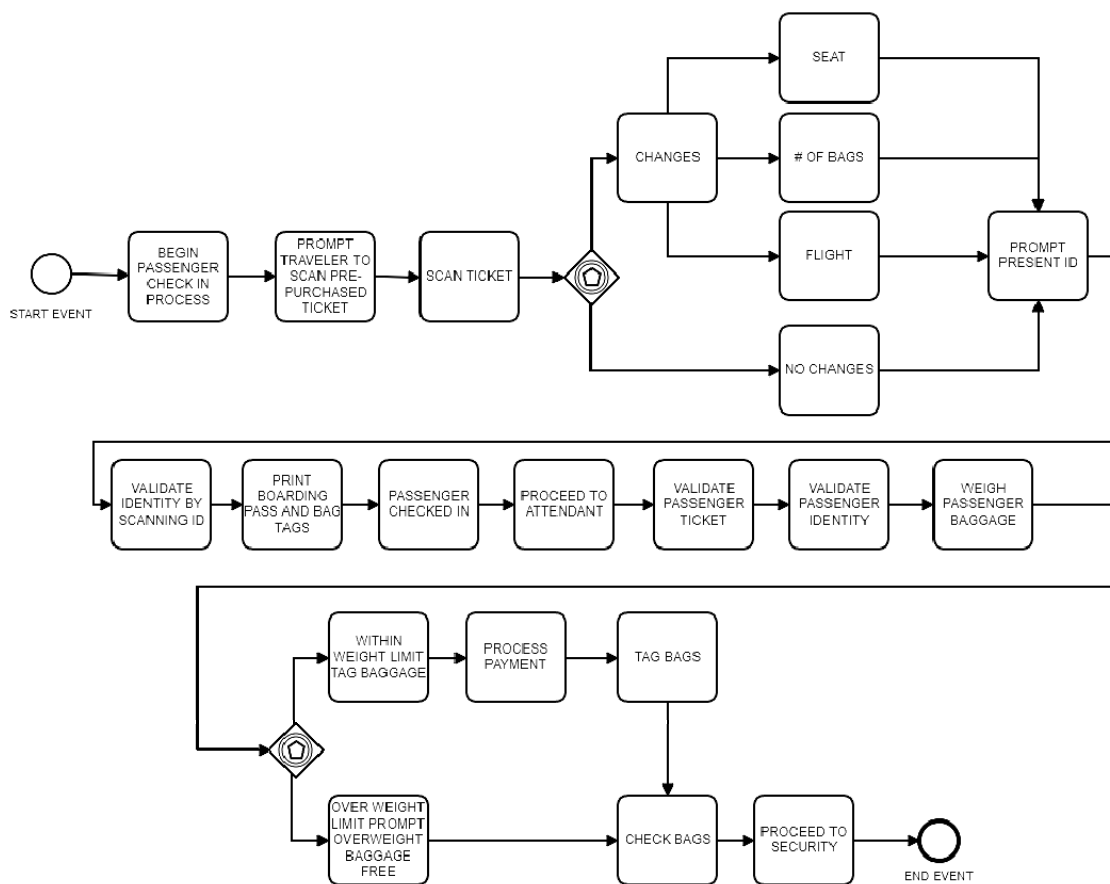
Figure 1. Boarding procedure model.

2.2. Resequencing Capability

It is essential to define the aim of each modeling transformation before determining the relevant BP model features. This paper paves the way for a transformation mechanism aiming to redesign the logic within the BP (i.e., the way it is executed). The BP behavior heuristic is a notion that deals with the execution order of activities and the way they are scheduled and assigned for execution [15]. This category is the most relevant to database-like optimization [8], and includes resequencing, parallelism and knockout parts. According to Reijers and Mansar [16], the parallelism and knockout heuristic are specific forms of the resequencing heuristic due to compulsory allocation of tasks to parallel branches and knockout parts in a growing order of effort and decreasing order of termination probability. A BP model should incorporate such features to be amenable to cost-based optimization solutions, and their assessment signifies the redesign capability, which is an integral check for the overall eligibility of the model.

In this paper, the authors focus on the resequencing heuristic. BPMN models lacking flexibility in task ordering are ineligible for data-centric workflow optimization [8]. Resequencing covers the optimization that involves changing the execution order of activities, while at the same time preserving the process semantics and correctness. This is performed in a cost-based manner rather than using ad-hoc heuristics. Due to the fact that there is no additional information or specification beyond the model depiction, the resequencing capability for the indicated example cannot be predetermined, and has to be either investigated or manually derived. In our case, the resequencing capability is speculative, as the investigation procedure is currently under development. Indicatively, it involves (a) listing the activities that can be resequenced; and (b) extracting dependency constraints that the resequencing should comply with in order to yield exactly the same output as before.
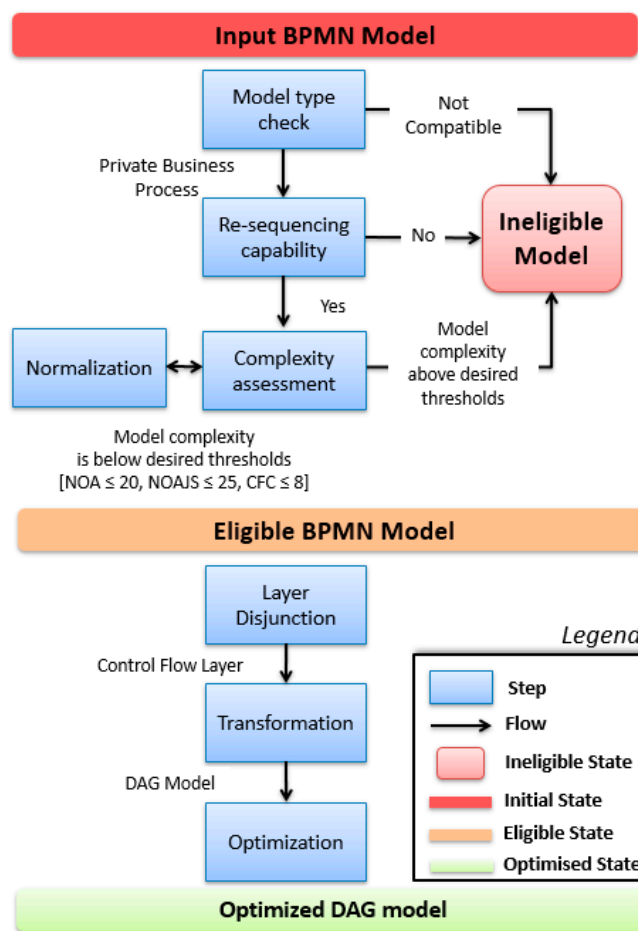
Figure 2. Eligibility assessment of BPMN models as an integral part of optimization.

2.3. Complexity Assessment

Evaluating and ultimately reducing BPMN model complexity improves, among others aspects, the correctness, maintainability, and understandability of BP models [17]. This fact has motivated researchers to propose various metrics in the area of BP measurement. Most of the initiatives are adaptations from the software engineering field, and many of them lack empirical validation [18]. A BPMN model complexity cannot be directly determined by only one type of metric [19]. Cardoso [20] identifies four main complexity perspectives: activity complexity, control-flow complexity, data-flow complexity, and resource complexity. In the current study, the authors have focused on complexity perspectives that affect the control flow of a BP and the number of control-flow elements, so that the input model can be disassembled to the control-flow layer, transformed, and further optimized. The selected metrics are: (a) number of activities (NOA); (b) number of activities, joins, and splits in a process (NOAJS) that are inspired by the lines-of-code (LOC) metric [21]; and (c) a control-flow complexity (CFC) metric [22] that calculates the complexity of XOR-split, OR-split, and AND-split constructs, and is inspired by McCabe's cyclomatic [23]. CFC analysis is useful in evaluating the difficulty of producing a BPMN process design before implementation, and by incorporating it in the process development cycle there is a considerable effect on the design phase, leading to increasingly optimized processes [24].

It should be noted that according to Cardoso [25], the CFC metric should not be used in isolation to effectively evaluate the overall complexity of a BP, because it only analyzes a process from the control-flow point of view. Similarly, the NOA and NOAJS metrics are useful and straightforward to calculate, but should accompany other complexity metrics (i.e., control flow in this paper). In the remainder of this work, we will assume that the mechanism filters out all complex processes;

more specifically, after the complexity assessment phase, it keeps models with NOA ≤ 20 and NOAJS ≤ 25 CFC ≤ 8, as it is at an early stage of development and experimentation. For establishing the complexity metric thresholds, the authors took into consideration the thresholds introduced in [26] for the evaluation of BP model modifiability. In this work, Yahya et al. define as moderate linguistic interpretation the thresholds 12 < NOA ≤ 26, 17 < NOAJS ≤ 33 and 3 < CFC ≤ 9. As most of the BP examples used in relevant literature fall within these boundaries, the authors selected a set of values that allow for a balance between eligible/ineligible BP models. This set is customizable and is aimed at progressively exposing the assessment mechanism of more complex BPMN models. Complexity assessment is interrelated with normalization, as shown in Figure 2, because the complexity of the initial model will most likely decrease after subjecting it to normalization. Yet, in many cases, and especially for the NOAJS metric, this does not betide, due to the fact that increasing structuredness by adding join constructs results in an increased value of NOAJS. The complexity metrics are calculated for the BPMN example. The NOA metric measures the number of activities (in this case tasks) and NOAJS the number of all activities, joins and splits of the process—in this case NOA (P) = 22 and NOAJS (P) = 24. The CFC metric evaluates the complexity of the XOR-split, OR-split, and AND-split constructs. The event-based gateways of the BPMN model are considered XOR gateways for measuring their complexity. The $CFC_{XOR-split}$, $CFC_{OR-split}$, and $CFC_{AND-split}$ functions are calculated as follows:

$CFC_{XOR-split}$ (1) = fan-out => $CFC_{XOR-split}$ (1) = 2 (event-based gateway for tracking changes)
$CFC_{XOR-split}$ (2) = fan-out => $CFC_{XOR-split}$ (2) = 2 (event-based gateway for weight limit)
$CFC_{OR-split}$ = 0
$CFC_{AND-split}$ = 0

The absolute control-flow complexity of the boarding procedure model is:

$$CFC_{abs}(P) = \left(\sum_{i \in (XOR-splits\ of\ P)} CFC_{XOR-split}i \right) + \left(\sum_{j \in (OR-splits\ of\ P)} CFC_{OR-split}j \right)$$
$$+ \left(\sum_{k \in (AND-splits\ of\ P)} CFC_{AND-split}k \right) => CFC_{abs}(P) = (2+2) + 0 + 0$$
$$=> CFC_{abs}(P) = 4$$

The boarding procedure initially had a complexity metric that exceeded the threshold set (NOA = 22). After normalization, the model is harmonized with the complexity requirements and is eligible for the next steps (transformation and optimization).

2.4. Normalization

Despite the fact that there are approaches that either provide guidelines on how to use the notation [26–28] or the background to analyze the modeling process itself [29,30], the majority of BPMN models do not adhere to these principles. The same semantics can be represented using various but behaviorally equivalent BPMN model structures. According to Yahya et al. [26], model quality improvement can be achieved through modeling guidelines, refactoring techniques, and transformation rules. To provide more-structured BP models, the authors considered established modeling recommendations [27,28], and employed the repository of equivalence patterns in BPMN models [31]. The modeling guidelines that the authors combined for the normalization step are listed as follows:

- use as few elements in the model as possible;
- minimize the routing paths per element;
- use one start and one end event;
- retain the model as structured as possible;
- avoid OR routing elements;

- use verb–object activity labels;
- decompose a model with more than 50 elements;
- avoid implicit splits and joins;
- provide tool support for proper model decomposition;
- omit the throwing message event;
- establish a centrally maintained glossary;
- provide tool support for linguistic checks during the modeling process.

Since the ultimate aim of the proposed approach is automated BP optimization, the authors intend to normalize BPMN models through a tool. Previous research on structuring and/or improving process models has been consolidated in tools like BPStruct [32], AProMoRe [33] and BP-Quality [26]. BPStruct seeks to structure a model to the maximum possible extent, and the resulting model is then said to be maximally structured. In future work, the authors intend to automate this step using BPStruct.

The aim is to improve the structure of BPMN models, reduce their complexity and facilitate the transformation phase. After the normalization step, the model complexity is re-evaluated based on NOA, NOAJS and CFC metrics to establish whether the input model complexity is within the desired thresholds. The eligibility assessment is complete and the input BPMN model is either deemed ineligible, or it meets all the eligibility criteria and advances to the following steps with the ultimate goal of optimization. For the particular example, the authors followed guidelines such as "model as structured as possible", "avoid implicit splits and joins", and "use as few elements in the model as possible", resulting in a more structured model (Figure 3). The guideline "avoid OR routing elements" could not be followed, as replacing the OR gateway with the behaviorally equivalent model structures (combining AND–XOR gateways) would result in a considerable increase of the CFC metric. Each complexity metric column features two values, one calculated prior and one after normalization (in the case it occurred). The updated values of the selected complexity metrics are:

NOA (PN) = 19, NOAJS (PN) = 25
$CFC_{XOR-split}$ (1) = fan-out => $CFC_{XOR-split}$ (1) = 2 (XOR gateway for tracking changes)
$CFC_{XOR-split}$ (2) = fan-out => $CFC_{XOR-split}$ (2) = 2 (XOR gateway for weight limit)
$CFC_{OR-split} = 2^{3-1}$ => $CFC_{OR-split} = 4$ (OR gateway for specifying changes)
$CFC_{AND-split} = 0$

$$CFC_{abs}(P_N) = \left(\sum_{i \in (XOR-splits\ of\ P)} CFC_{XOR-split}i\right) + \left(\sum_{j \in (OR-splits\ of\ P)} CFC_{OR-split}j\right)$$
$$+ \left(\sum_{k \in (AND-splits\ of\ P)} CFC_{AND-split}k\right) => CFC_{abs}(P_N) = (2+2) + 2^{3-1} + 0$$
$$=> CFC_{abs}(P_N) = 8$$

It is important to mention the increase of the NOAJS metric after the normalization step, due to the insertion of join constructs for improving the degree of structuredness. According to Dumas et al. [32], a well-known property of process models is that of block-structuredness, meaning that for every node with multiple outgoing arcs (a split) there is a corresponding node with multiple incoming arcs (a join) such that the subgraph between the split and the join forms a single-entry/single-exit (SESE) region. Inversely, by decreasing the NOAJS metric of a model, it is not necessarily converted to a more comprehensible and modifiable model if this is achieved by eliminating join constructs. The increase of the CFC metric is attributed to an implicit gateway in the initial model that was converted to an OR gateway.

The following steps 5, 6, and 7 are out of the scope of this paper; the authors chose to include them here to better demonstrate the end-to-end approach that is under development.
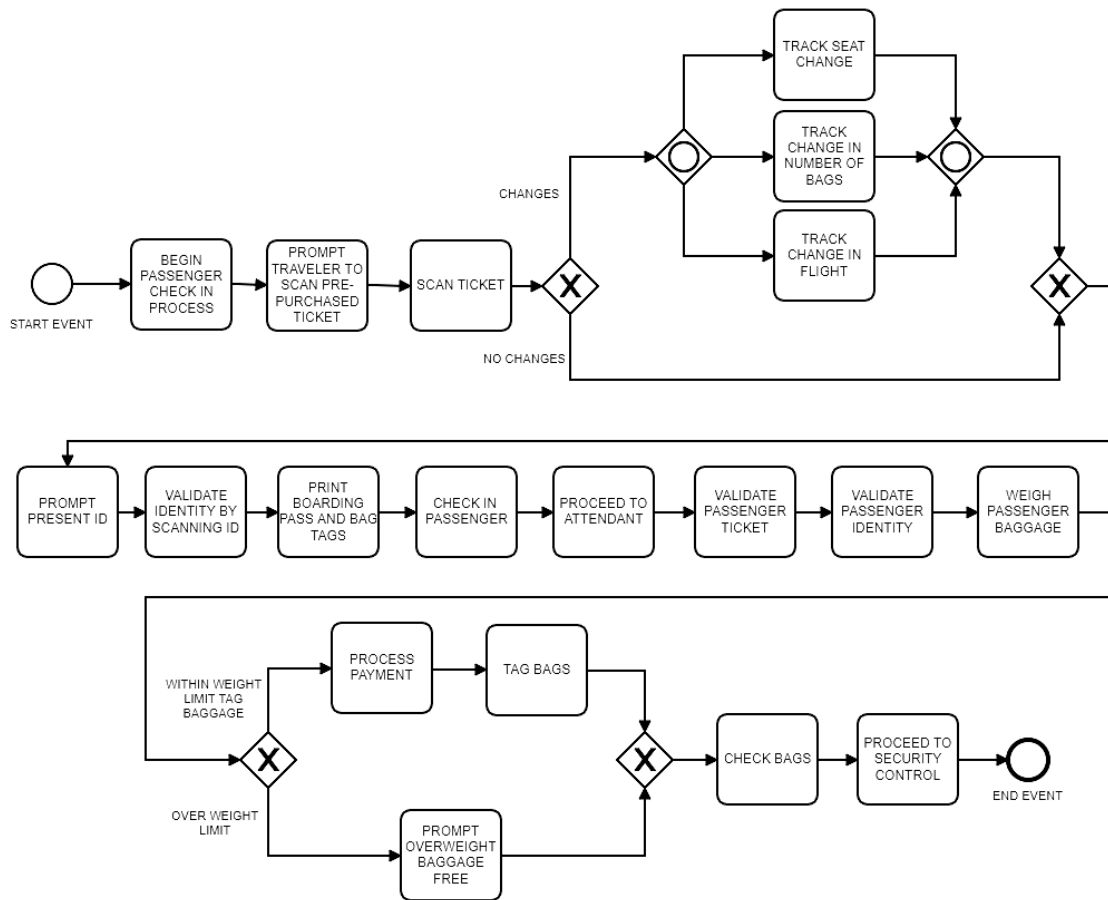
Figure 3. Boarding procedure model after normalization.

2.5. Layer Disjunction

A BPMN model is assembled of elements belonging to the five fundamental categories (flow objects, connecting objects, data, swimlanes and artifacts). Depending on their usability, these are subsequently allocated to five layers (i.e., control-flow, data, readability, organizing and execution layers). The optimization framework aims to optimize: (i) the execution performance, and (ii) the functional aspects that are represented through flow objects and a subset of connecting objects (i.e., sequence and message flows). As a result, the transformation process will focus on the control-flow layer of the model serving as input.

2.6. Transformation

The transformation phase will match different aspects of BPs and data-intensive workflows by creating a paradigm that associates a BP model with a data-flow execution plan. To accomplish this, the notion of artificial tasks termed as dummy tasks is introduced in [8]. Overall, the combination of normal and dummy tasks with appropriately set statistical metadata will enable modeling the token flow in BPs and pave the way for performance optimization. The principal assumption is that the operational behavior of a BP modeled in a BPMN diagram can be emulated by an equivalent data-flow model, despite the fact that they stand for different entities. Following the prior disjoining procedure, this step will consist of the BPMN-to-DAG transformation process and its subprocesses which are currently under development. Each BPMN element will be transformed in accordance to the proposed symbol mapping [8]. It is important to clarify that this transformation paradigm also supports process models with loops (e.g., when gateways are accompanied by cycles). In these cases, the combination of the proposed symbol mapping in [8] regarding loops and gateways renders the graph acyclic.

2.7. Optimization

The last step takes the initial DAG and the set of statistical and dependency metadata and produces an optimal execution plan. The explicit optimization metric used is the sum of the average execution times of each task that correspond to the product of the task cost and selectivity values of the preceding tasks. In an aggregate manner, this sum defines the average running (cycle) time of the process.

3. Examples of Assessing BPMN Input Models

The eligibility mechanism is further evaluated for cases extracted from the SOA-based business process database [13].

The SOA-based business process database contains BPs of organizations operating in different sectors modeled using BPMN. The aim is to showcase the eligibility criteria that an input model must satisfy and also demonstrate the range of ineligible cases that can occur. The evaluation intends to manifest the positive effects of normalization through the selected complexity metrics. The authors selected 10 indicative process models that cover a wide range of scenarios and evaluated them with the proposed mechanism, as shown in Table 1. Each line denotes the selected business process example, while columns present the necessary information for their designation and eligibility assessment. For each example, the authors recorded the descriptive label of the process, model type, resequencing capability and calculated complexity metrics. These result in the overall model eligibility column (eligible/ineligible) and a column for tracking the unmet criteria for the ineligible cases. The complexity metrics are calculated once in structured models and twice in unstructured ones (two entries separated by virgule), to better demonstrate the effects of normalization. Due to the fact that there is no additional information or specification beyond the models' depictions and the resequencing capability is not explicitly mentioned, the latter is manually derived.

Table 1. BPMN examples from the SOA-based business process database [13].

No.	Business Process	Model Type	Resequencing	NOA	NOAJS	CFC	Eligibility	Criteria That Are Not Satisfied
1	Online Purchase (No. 139)	Public Process	Not Explicit	9/-	12/-	8/-	Ineligible	• Incorrect model type • Complexity above set thresholds
2	VIP Customer Problem (No. 62)	Public Process	Not Explicit	12/-	17/-	8/-	Ineligible	• Incorrect model type • Complexity above set thresholds
3	Product Marketing Plan (No. 135)	Private Process	Not Explicit	37/-	47/-	7/-	Ineligible	• Complexity above set thresholds
4	Employee Recruitment (No. 197)	Private Process	Not Explicit	25/-	37/-	15/-	Ineligible	• Complexity above set thresholds
5	IT Help Desk (No. 195)	Private Process	Not Explicit	17/-	22/-	9/-	Ineligible	• Complexity above set thresholds
6	Hardware Retail (No. 17)	Private Process	Not Explicit	8/-	14/-	5/-	Eligible	-
7	Charity (No. 53)	Private Process	Not Explicit	6/-	9/-	4/-	Eligible	-
8	Account Opening (No. 122)	Private Process	Not Explicit	9/-	13/-	4/-	Eligible	-
9	Customized PC Purchase (No. 196)	Private Process	Not Explicit	10/-	11/-	2/-	Eligible	-
10	Boarding Procedure (No. 198)	Private Process	Not Explicit	22/19	24/25	4/8	Eligible	-

The first two examples depict two public processes that, according to the first criterion, are not of a compatible model type. The next three BPMN models are of compatible model types (private non-executable process) but exceed the complexity metric thresholds. In particular, the "Product Marketing Plan" model consists of an increasing number of activities, joins and splits (NOA = 37, NOAJS = 47) with acceptable control-flow complexity (CFC = 7). The "Employee Recruitment" model depicts a model with high values of all measured complexity metrics (i.e., NOA = 25, NOAJS = 37, CFC = 15). The "IT Help Desk" model is also rejected, due to a control-flow metric (CFC = 9) higher than the set threshold. The remaining five examples are private processes with complexity metrics within the acceptable limits, and are characterized as eligible. The "Hardware Retail", "Charity", "Account Opening", "Customized PC Purchase", and "Boarding Procedure" models have low complexity values, which demonstrates their relatively simple structures and limited sizes. Four out of five eligible models (except for "Boarding Procedure") do not need normalization, as they adhere to the modeling guidelines of the normalization step and also feature the block-structuredness property (i.e., every subgraph between the split and the join forms a single-entry/single-exit (SESE) region [32]).

The eligibility mechanism is also evaluated using BPMN models located in relevant research papers, as shown in Table 2. A major difference with the SOA repository is that the resequencing criterion is explicitly stated in the selected models based on their attributes. In cases where the capability is not explicitly stated in the corresponding paper, the authors assume that there is no flexibility in task ordering and the model is considered ineligible. The first model depicts a "Healthcare Scenario" [34] public process that it is not the desired input model type. The following two examples ("Loan Application" [35] and "Maintenance" [36]) do not mention resequencing as an explicit capability, a fact that classifies them as ineligible despite their complexity metrics being within limits. The "Evaluate Quote Process" [37] and "Medical Assessment/Treatment" [38] models are rejected due to a control-flow complexity (CFC = 9) higher than the set threshold.

Table 2. BPMN examples in relevant research papers.

No.	Business Process	Model Type	Resequencing	NOA	NOAJS	CFC	Eligibility	Criteria That Are Not Satisfied
1	Healthcare Scenario [34]	Public Process	No/not explicit	14/-	20/-	6/-	Ineligible	• Incorrect model type • Not explicit resequencing
2	Loan Application [35]	Private Process	No/not explicit	5/-	7/-	4/-	Ineligible	• Not explicit resequencing
3	Maintenance [36]	Private Process	No/not explicit	12/-	16/-	5/-	Ineligible	• Not explicit resequencing
4	Evaluate Quote Process [37]	Private Process	Yes	7/-	13/-	9/-	Ineligible	• Complexity above set thresholds
5	Medical Assessment/ Treatment [38]	Private Process	Yes	11/-	18/-	9/-	Ineligible	• Complexity above set thresholds
6	Car-Rental [39]	Private Process	Yes	10/-	13/-	3/-	Eligible	-
7	Property Valuation [40]	Private Process	Yes	16/-	20/-	5/-	Eligible	-
8	Admission Process [41]	Private Process	Yes	16/-	19/-	6/-	Eligible	-
9	Emergency Ward of a Hospital [42]	Private Process	Yes	10 /10	19/19	9/7	Eligible	-
10	Client Complaint for Product Defect [26]	Private Process	Yes	15 /14	28/18	20/8	Eligible	-

The next three models ("Car-Rental" [39], "Property Valuation" [40] and "Admission Process" [41]) feature eligible models as they are private non-executable processes with explicit resequencing

capabilities and metric values within the desired thresholds. The next model, "Emergency Ward of a Hospital" [42], has an unacceptable CFC metric (CFC = 9). Following the normalization guidelines in [31], the authors replaced a loop modeled using control flow with an equivalent loop activity that preserved semantic correctness and simultaneously reduced the CFC metric to an acceptable value (CFC = 7). The last model ("Client Complaint for Product Defect" [26]) depicts a process with resequencing capability and increased initial complexity metrics (NOA = 15, NOAJS = 28, CFC = 20). In [26], a guiding framework (BP-Quality) that supports designers in improving the quality of BP models was used in the model. This resulted in noticeably decreased complexity metrics (NOA = 14, NOAJS = 18, CFC = 8) below the set threshold, thus the model is eligible. Similarly, to the eligible models of the first set (Table 1), three out of five eligible models do not need normalization, as they adhere to the modeling guidelines of the normalization step and also feature the block-structuredness property. The remaining two were normalized and the complexity metrics decreased.

4. Directions and Future Work

This paper introduced a comprehensive assessment mechanism that measures the eligibility of a BPMN model (i.e., its capability to be used as input in a redesign initiative), in our case performance optimization using data-centric workflow optimization algorithms. In terms of comprehensibility, this work provides a better understanding of the BPMN model through evaluation of the BP based on its structural complexity and normalization, aiming for more structured diagrams with decreased semantic complexity.

As illustrated in Figure 2 and currently under investigation by the authors, an eligible model is fit for further processing that involves: (a) layer disjunction to generate the BP control flow, (b) transformation to a DAG, and (c) optimization to produce an optimal execution plan. One of the main contributions of this work is that the eligibility mechanism is not restricted to this redesign initiative, as it can plug in different disciplines depending on the particular perspective (decoded to redesign heuristics). It is necessary to denote that a BP model should incorporate such features to be amenable to cost-based optimization solutions, and their assessment signifies the redesign capability—an essential criterion for the overall eligibility of the BP model. The proposed mechanism was evaluated on two sets of BP models extracted from the SOA-based BPD repository and relevant research papers. The selection of the BPMN models as potential inputs showcased the range of possible outcomes regarding their eligibility and complexity assessment. On this basis, three models that featured an unacceptable model type (public processes) and two that did not have an explicit resequencing capability have been deemed ineligible for optimization. For the remaining 15 BPMN models, the complexity metrics were calculated, and five had values above the desired thresholds. The 10 BPMN models that met the criteria are eligible models for transformation and optimization. The next step involved the assessment of their structuredness, in which seven of them were substantially structured and three required normalization. A motivating example was selected and the practice of passing every step of the eligibility mechanism was demonstrated. The structured and semantically equivalent model was transformed to a DAG to showcase the aspiration and applicability of the mechanism towards transformation.

The main open research issues of the proposed assessment mechanism are: (a) the lack of a concise analysis of potential redesign heuristics and their implementation, (b) the derivation of a suitable combinatorial complexity metric with clearly defined thresholds, and (c) the mapping of optimized execution plans back to BPMN models after the optimization step. Future work will be motivated by the absence of adequate quantitative support for many redesign heuristics and methodological approaches for "measuring" the capability of a BP model to optimally change. There is a research gap in the literature regarding the identification, documentation, and ultimately a suggestion of BPR practices prior to their application. The authors are currently working on establishing quantitative measures regarding the assessment of redesign features in a BPMN model required for the application of BPR practices.

The authors are also working on a comprehensive complexity metric that combines a number of established initiatives reflecting the number and size of control constructs and process intricacy. The selection of metrics is based on literature findings and particular criteria related to transformation and redesign applicability. Through extensive testing of the multi-facet metric using BP repositories, the definition of practical thresholds will assist users in selecting suitable BPs for optimization. The next research steps involve the introduction and validation of a systematic set of rules in the form of a BPMN-to-DAG transformation process, through the testing of various eligible BPMN models. The extracted DAG models will be experimented using data-centric workflow optimization techniques. Due to the need for end-to-end solutions, the authors are also working on mapping the optimized execution plan back to a BPMN model. This would ideally be exposed as a software plugin to existing platforms and encompass all the mapping and optimization procedures. An extension of this work is the examination of dependency-aware optimization algorithms that consider parallelism constraints, capability information and blocking versus pipelining information, in an effort to provide a solid and comprehensive optimization perspective.

Author Contributions: Conceptualization, G.T., K.V. and A.G.; Data curation, G.K.; Investigation, K.V.; Methodology, G.T. and A.G.; Resources, G.K.; Supervision, K.V. and A.G.; Validation, G.K.; Writing—original draft, G.T.; Writing—review & editing, K.V. and A.G.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H.A. *Fundamentals of Business Process Management*; Springer: Berlin/Heidelberg, Germany, 2013.
2. Abdi, N.; Zarei, B.; Vaisy, J.; Parvin, B. Innovation models and business process redesign. *Int. Bus. Manag.* **2011**, *3*, 147–152.
3. Buzacott, J.A. Commonalities in reengineered business processes: models and issues. *Manag. Sci.* **1996**, *42*, 768–782. [CrossRef]
4. La Rosa, M.; Dumas, M.; Ter Hofstede, A.H.; Mendling, J. Configurable multi-perspective business process models. *Inf. Syst.* **2011**, *36*, 313–340. [CrossRef]
5. Van der Aalst, W.M. Re-engineering knock-out processes. *Decis. Support Syst.* **2001**, *30*, 451–468. [CrossRef]
6. Kumar, A.; Indradat, P. Optimizing process model redesign. In Proceedings of the International Conference on Service-Oriented Computing, Banff, AB, Canada, 10–13 October 2016; pp. 39–54.
7. Doomun, R.; Vunka Jungum, N. Business process modelling, simulation and reengineering: call centres. *Bus. Process Manag. J.* **2008**, *14*, 838–848. [CrossRef]
8. Gounaris, A. Towards automated performance optimization of BPMN business processes. In Proceedings of the East European Conference on Advances in Databases and Information Systems, Prague, Czech Republic, 28–31 August 2016; pp. 19–28.
9. Kougka, G.; Gounaris, A.; Simitsis, A. The Many Faces of Data-centric Workflow Optimization: A Survey. *arXiv* **2017**, arXiv:1701.07723. [CrossRef]
10. Palma-Mendoza, J.A.; Neailey, K.; Roy, R. Business process re-design methodology to support supply chain integration. *Int. J. Inf. Manag.* **2014**, *34*, 167–176. [CrossRef]
11. Röder, A.; Tibken, B. A methodology for modeling inter-company supply chains and for evaluating a method of integrated product and process documentation. *Eur. J. Oper. Res.* **2006**, *169*, 1010–1029. [CrossRef]
12. Wiśniewski, P. Decomposition of business process models into reusable sub-diagrams. In Proceedings of the ITM Web of Conferences, EDP Sciences, Lublin, Poland, 23–25 November 2017; Volume 15, p. 01002.
13. BPOSCTeam SOA-Based Business Process Database. Available online: https://sites.google.com/site/bposcteam2015/ressources (accessed on 11 March 2019).
14. Object Management Group (OMG) Business Process Model and Notation (BPMN) Version 2.0 2011. Available online: http://www.oatsolutions.com.br/artigos/SpecBPMN_v2.pdf (accessed on 30 June 2019).
15. Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H.A. *Fundamentals of Business Process Management*; Springer: Berlin/Heidelberg, Germany, 2018; ISBN 978-3-662-56509-4.

16.　Reijers, H.A.; Mansar, S.L. Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. *Omega* **2005**, *33*, 283–306. [CrossRef]

17.　Cardoso, J.; Mendling, J.; Neumann, G.; Reijers, H.A. A discourse on complexity of process models. In Proceedings of the International Conference on Business Process Management, Vienna, Austria, 4–7 September 2006; pp. 117–128.

18.　Kluza, K.; Nalepa, G.J. Proposal of square metrics for measuring business process model complexity. In Proceedings of the 2012 Federated Conference on Computer Science and Information Systems (FedCSIS), Wroclaw, Poland, 9–12 September 2012; pp. 919–922.

19.　Mendling, J. Metrics for business process models. In *Metrics for Process Models*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 103–133.

20.　Cardoso, J. Business process control-flow complexity: Metric, evaluation, and validation. *Int. J. Web Serv. Res.* **2008**, *5*, 49–76. [CrossRef]

21.　Jones, C. *Programming Productivity*, 1st ed.; McGraw-Hill College: New York, NY, USA, 1986; ISBN 978-0-07-032811-2.

22.　Cardoso, J. How to measure the control-flow complexity of web processes and workflows. *Work. Handb.* **2005**, *2005*, 199–212.

23.　McCabe, T.J. A complexity measure. *IEEE Trans. Softw. Eng.* **1976**, *4*, 308–320. [CrossRef]

24.　Rolón, E.; Cardoso, J.; García, F.; Ruiz, F.; Piattini, M. Analysis and validation of control-flow complexity measures with bpmn process models. In *Enterprise, Business-Process and Information Systems Modeling*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 58–70.

25.　Cardoso, J. Control-flow complexity measurement of processes and Weyuker's properties. In Proceedings of the 6th International Enformatika Conference, Budapest, Hungary, 26–28 October 2005; Volume 8, pp. 213–218.

26.　Yahya, F.; Boukadi, K.; Ben-Abdallah, H. Improving the quality of Business Process Models: Lesson learned from the State of the Art. *Bus. Process Manag. J.* **2018**. [CrossRef]

27.　Mendling, J.; Reijers, H.A.; van der Aalst, W.M. Seven process modeling guidelines (7PMG). *Inf. Softw. Technol.* **2010**, *52*, 127–136. [CrossRef]

28.　Leopold, H.; Mendling, J.; Günther, O. Learning from quality issues of BPMN models from industry. *IEEE Softw.* **2016**, *33*, 26–33. [CrossRef]

29.　Pinggera, J.; Zugal, S.; Weidlich, M.; Fahland, D.; Weber, B.; Mendling, J.; Reijers, H.A. Tracing the process of process modeling with modeling phase diagrams. In Proceedings of the International Conference on Business Process Management, Clermont-Ferrand, France, 29 August 2011; pp. 370–382.

30.　Claes, J.; Vanderfeesten, I.; Reijers, H.A.; Pinggera, J.; Weidlich, M.; Zugal, S.; Fahland, D.; Weber, B.; Mendling, J.; Poels, G. Tying process model quality to the modeling process: the impact of structuring, movement, and speed. In Proceedings of the International Conference on Business Process Management, Tallinn, Estonia, 3–6 September 2012; pp. 33–48.

31.　Kluza, K.; Kaczor, K. Overview of BPMN model equivalences: towards normalization of BPMN diagrams. In Proceedings of the 8th Workshop on Knowledge Engineering and Software Engineering (KESE2012) at the at the Biennial European Conference on Artificial Intelligence (ECAI 2012), Montpellier, France, 28 August 2012; Volume 28, pp. 38–45.

32.　Dumas, M.; La Rosa, M.; Mendling, J.; Mäesalu, R.; Reijers, H.A.; Semenenko, N. Understanding business process models: the costs and benefits of structuredness. In Proceedings of the International Conference on Advanced Information Systems Engineering, Gdansk, Poland, 25–29 June 2012; pp. 31–46.

33.　La Rosa, M.; Reijers, H.A.; Van Der Aalst, W.M.; Dijkman, R.M.; Mendling, J.; Dumas, M.; García-Bañuelos, L. APROMORE: An advanced process model repository. *Expert Syst. Appl.* **2011**, *38*, 7029–7040. [CrossRef]

34.　Knuplesch, D.; Reichert, M.; Fdhila, W.; Rinderle-Ma, S. On enabling compliance of cross-organizational business processes. In *Business Process Management*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 146–154.

35.　Rogge-Solti, A.; Weske, M. Prediction of business process durations using non-Markovian stochastic Petri nets. *Inf. Syst.* **2015**, *54*, 1–14. [CrossRef]

36.　Kannengiesser, U. Can we engineer better process models? In Proceedings of the 17th International Conference on Engineering Design, Palo Alto, CA, USA, 24–27 August 2009; pp. 527–538.

37. Kolar, J.; Dockal, L.; Pitner, T. A Dynamic Approach to Process Design: A Pattern for Extending the Flexibility of Process Models. In Proceedings of the IFIP Working Conference on The Practice of Enterprise Modeling, Riga, Latvia, 6–7 November 2013; pp. 176–190.

38. Herbert, L.T.; Sharp, R.; Hansen, M.R. *Specification, Verification and Optimisation of Business Processes: A Unified Framework*; Technical University of Denmark: Lyngby, Denmark, 2014.

39. Truong, T.-M.; Lê, L.-S. Towards a Formal Framework for Business Process Re-Design Based on Data Mining. In *Enterprise, Business-Process and Information Systems Modeling*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 250–265.

40. Kannengiesser, U.; Zhu, L. Towards Concise Architectures for Flexible Business Processes. *Adv. Eng. Inform.* **2011**, 61–75.

41. Allah Bukhsh, Z.; van Sinderen, M.J.; Sikkel, N.; Quartel, D. Understanding modeling requirements of unstructured business processes. In Proceedings of the 14th International Joint Conference on e-Business and Telecommunications, ICETE 2017, Madrid, Spain, 26–28 July 2017.

42. Mannhardt, F.; de Leoni, M.; Reijers, H.A.; van der Aalst, W.M. Data-driven process discovery-revealing conditional infrequent behavior from event logs. In Proceedings of the International Conference on Advanced Information Systems Engineering, Essen, Germany, 12–16 June 2017; pp. 545–560.