

Evaluating data-centric process approaches: Does the human factor factor in?

Hajo A. Reijers¹ · Irene Vanderfeesten² · Marijn G. A. Plomp⁵ · Pieter Van Gorp² · Dirk Fahland² · Wim L. M. van der Crommert³ · H. Daniel Diaz Garcia⁴

Received: 30 September 2014 / Revised: 10 August 2015 / Accepted: 12 August 2015 / Published online: 9 February 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The Business Process Management field addresses design, improvement, management, support, and execution of business processes. In doing so, we argue that it focuses more on developing modeling notations and process design approaches than on the needs and preferences of the individual who is modeling (i.e., the user). New data-centric process modeling approaches are taken as a relevant and timely stream of process design approaches to test our argument. First, we provide a review of existing data-centric process approaches, culminating in a theoretical classification framework. Next, we empirically evaluate three specific approaches with regard to the claims they make. We had participants representative of actual users try out these approaches on realistic scenarios via a series of workshops. Participants assessed to what extent quality claims from the literature could be recognized within the workshop sessions. The results of this evaluation substantiate a number

of claims behind the approaches, but also identify opportunities to further improve them. Most prominently, we found that the usability aspects of all considered approaches are a source of concern. This leads us to the insight that usability aspects of process design approaches are crucial and, in the perception of groups representative of actual users, leave much to be desired. In that sense, our research can be seen as a wake-up call for process modeling notation designers to consider the usability side—and as such, the interest of the human modeler—more than is currently the case.

Keywords Process modeling · User · Usability · Data-centric · Evaluation · Review

1 Introduction

The Business Process Management (BPM) field addresses the design, improvement, management, support, and execution of business processes. Its cornerstone artifact is the process model, which is used to describe the activities to be executed to handle a case, as well as their order of execution. There is a range of modeling approaches available, with two main categories of these being the more traditional, activity-oriented process models and the more recent, data-centric approaches. We observe that in general, the emphasis in the BPM community seems to be on developing modeling notations and process design approaches without paying much attention to the needs and preferences of the individual who is modeling. In particular, while claims about the high usability of new methods and notations are certainly being made, it requires the actual involvement of people representative of their users to substantiate such claims. This is the starting point for our current research, where we evaluate three data-centric process modeling approaches with regard to these claims.

Communicated by Dr. Selmin Nurcan.

✉ Hajo A. Reijers
h.a.reijers@vu.nl

Marijn G. A. Plomp
m.g.a.plomp@vu.nl

¹ Faculty of Sciences, Information Management & Software Engineering, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

² Eindhoven University of Technology, Eindhoven, The Netherlands

³ Precedence, Maastricht, The Netherlands

⁴ Ericsson, Mexico City, Mexico

⁵ Faculty of Economics and Business Administration, Knowledge, Information and Innovation (KIN) Research Group, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

In the last decade, a number of data-centric approaches have been developed to extend or counter the overly activity-oriented way of designing and executing business processes, as activity-oriented process models had become known for a number of limitations. For example, in dynamic environments, they can become too rigid, which makes it difficult to cope with changes in the process. The central idea behind data-centric approaches is that data objects/elements/artifacts can be used to enhance a process-oriented design or even to serve as the fundament for such a design. This has certain advantages, varying from increasing flexibility in process execution and improving reusability to actually being able to capture processes where data play a relevant role. We refer the reader who wants to get a quick yet thorough understanding of such data-centric approaches to the online tutorials on three of these at: <http://is.ieis.tue.nl/staff/ivanderfeesten/Research/Data-centric/>. Despite considerable research efforts, it seems fair to say that data-centric process approaches have not become mainstream at this point. Yet, the recurrence of new variants that treat data as a first-class citizen suggests that the core idea in itself has merit. Therefore, we consider data-centric process modeling approaches as a relevant and timely stream of process design approaches to further examine with regard to the level to which they live up to their promises.

In this paper, we start with a review of existing data-centric process approaches and determine what conceptual similarities and differences exist between them. This forms the foundation for the main question that interests us in this paper: How good are these various data-centric approaches actually according to their users? We focus on the *perceived* quality of the data-centric process modeling approaches. It is far from clear—for these newly developed modeling techniques—how their usability claims relate to the way people using these methods experience their usability. An answer to our main question provides insights to the developers of the approaches to understand to what extent their claims are recognized and appreciated by prospective users.

There are previous studies in which existing data-centric process approaches were compared. Künzle and Reichert [10] have made a feature comparison of several approaches (including artifact-centric process models and case handling). Henriques and Silva [7] analyzed the support of data-related requirements like access and granularity for a number of object-oriented approaches. Our paper extends these studies in three respects. First, we are particularly interested in the degree to which these process approaches take human factors (i.e., the usability) into account. Second, our review of data-centric process approaches widens the scope of approaches under consideration, while deepening their conceptual understanding. The respective contribution is a classification of 14 data-centric process approaches. The classification shows how each approach is positioned

and through which interaction between process and data. Third, our work adds an empirical dimension by incorporating an evaluation of the actual use of three data-centric process approaches, namely Data-Driven Process Structures (DDPS), Product-Based Workflow Design (PBWD), and Artifact-Centric Process Models (ACPM). Our contribution here is an insight into the level to which the approaches live up to their promises, as well as some improvement opportunities for each of these.

The structure of this paper is as follows. Section 2 presents our review of data-centric process approaches, based on an extensive search of the literature and a classification of the identified approaches. Section 3 explains the methodology that has been followed to evaluate the quality—in particular the usability—of these methods. A discussion of the evaluation results can be found in Sect. 4. In Sect. 5, we conclude this paper with a summary, limitations of the study, and an outlook into future work.

2 Review

2.1 Identification

To identify the relevant literature on data-centric process approaches, we followed an approach originally proposed by Cooper [4] and applied in the BPM domain by Bandara [2]. We selected two subdomains for our search of the relevant literature: information systems (IS) and BPM. Within these subdomains, academic journals and conferences were selected as the primary sources of information. We identified the body of existing data-centric process modeling techniques by searching for articles that contained the following keywords (in the entire article text): *business process modeling approaches*, *business process modeling paradigms*, and *business process modeling methods*. The search results were refined using the following keywords: *information centric process*, *data centric process*, *data driven process*, *artifact centric process*, *product based process*, *document driven process*, and *object centric process*. The same keywords were used interchanging the word process for workflow. Additionally, we followed citations in the identified articles (“snowballing”) and included articles referred to us by researchers in the domain.

This literature search revealed 14 distinct approaches, each being described in several articles. We identified for each technique the two most cited articles using Google Scholar. One of these papers is typically a seminal paper on the approach. The other is either an introductory conference article or a follow-up article, extending the approach. As founding year of the approach, we took the publication year of the oldest paper. We also identified the parties proposing the approach. A chronologically ordered overview is given in Table 1.

Table 1 Data-centric process approaches

| Data-centric process approach | Proposers | Years |
|--|--|-------|
| WASA ₂ | University of Münster | 1999 |
| TriGSflow | University of Linz | 2000 |
| Object coordination nets | Westfälische Wilhelms-Universität, Universität-Gesamthochschule-Paderborn | 2001 |
| XDoc-WFMS | IIIT Hyderabad | 2002 |
| Product-Based Workflow Design (PBWD) | Eindhoven University of Technology, Deloitte | 2003 |
| Artifact-Centric Process Models (ACPM) | IBM Watson Research Laboratory | 2003 |
| XFlows | Consiglio Nazionale delle Ricerche IIT | 2005 |
| Document-driven workflows | Pennsylvania State University | 2005 |
| Case handling | Eindhoven University of Technology, Perceptive Software | 2005 |
| Data-Driven Process Structures (DDPS) | University of Ulm, University of Twente, Daimler Chrysler AG | 2007 |
| Object lifecycle compliant processes | IBM Zurich Research Laboratory, University of Zurich | 2007 |
| Object-centric process modeling | Queensland University of Technology, FlowConnect | 2010 |
| Product-Based Workflow Support (PBWS) | Eindhoven University of Technology | 2011 |
| Philharmonic Flows | University of Ulm, Persis GmbH | 2011 |

2.2 Classification

Next, we classified the identified data-centric process modeling techniques using the two most cited articles for each paper. Previous surveys [7, 10] classified techniques by their *features*. While using their views as input for our classification, we had a slightly different aim. We wished to identify a single, encompassing feature that would provide the essence of the approach with respect to the interaction between data and process. Also, since a technique is always developed for a specific purpose, we aimed to identify the *drivers* that motivated the particular development of an approach. These drivers mostly take on the form of a particular advantage that the approach brings or a specific problem that can be overcome.

To support our classification, we extracted the wealth of information from the articles in a uniform way using *reference concept maps* (RCMs) [22]. RCMs are essentially a graphical representation of concepts (represented as nodes labeled with the concept name) and the meaning given to these concepts in a paper (represented as edges labeled with verbs or expressions and the originating paper). Particularly, valuable information was often found in the related work sections in the consulted papers, in which references to related work could be found and, sometimes, explicit comparisons with other methods.

We kept on iterating in our search for a main distinguishing feature and the types of drivers until we arrived at a number between 5 and 7 for each of these two dimensions, a level of detail for a classification that is on the one hand

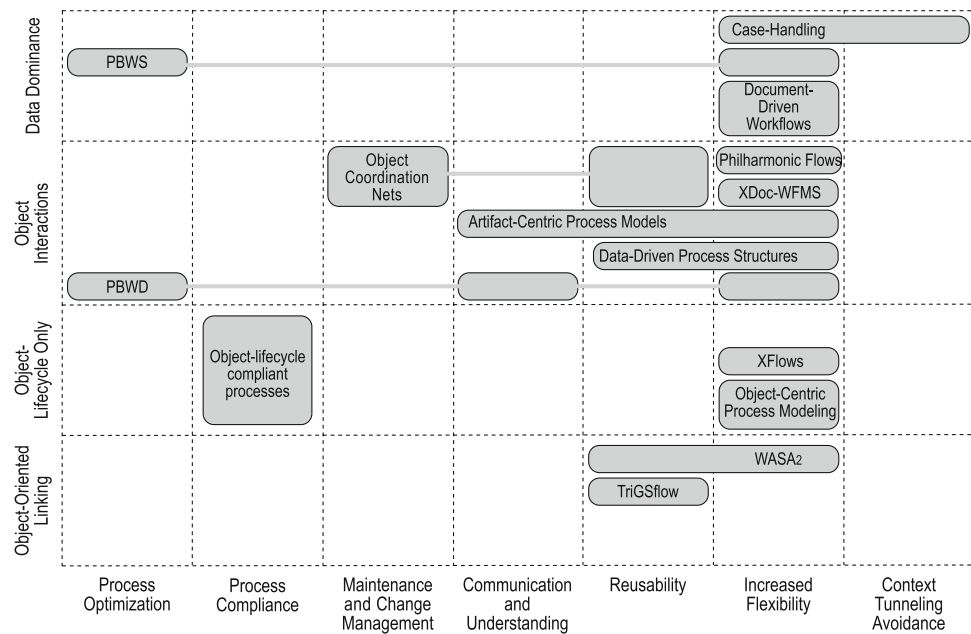
potentially insightful and on the other still manageable. In Fig. 1, all 14 approaches can be seen against the backdrop of (a) seven different drivers for data-centric process modeling approaches and (b) four different levels of integration of data in the approach. While each approach can be characterized by one or more drivers, each approach sits at one specific level of data integration. In other words, the drivers have a categorical, non-exclusive nature; the data integration levels are on an ordinal scale. The positioning of approaches within a cell is purely for graphical convenience and has no deeper meaning.

2.2.1 Drivers for developing data-centric process approaches

The set of main drivers that we have extracted from the literature are the following: (1) enabling *process optimization*, (2) enabling *process compliance*, (3) simplifying *maintenance and change management*, (4) supporting *communication and understanding*, (5) fostering *reusability*, (6) *increased flexibility*, and (7) *context tunneling avoidance*.

What can be seen in Fig. 1 is that the most frequent driver for the development of the data-centric process approaches under consideration is to increase the flexibility of the process: In 11 out of 14 of the approaches, this driver is explicitly mentioned. It also makes sense intuitively when considering that by giving the lifecycle of data more prominence in a process it is not necessary to completely specify and—in this way—straightjacket it.

Fig. 1 Classification of data-centric process approaches



The second most frequent driver is to allow more reuse of process knowledge, which is covered by 5 of the 14 approaches. The relation with data centrality in a process approach is that by making process specifications less explicit, they become more generic and therefore more easily applicable in other settings.

Each of the remaining drivers is tied to just one or two approaches each: process optimization to *PBWS* [24] and *PBWD* [21], process compliance to *object lifecycle compliant processes* [11], maintenance and change management to *object coordination nets* [27], communication and understanding to *ACPM* [16] and *PBWD*, and context tunneling avoidance to *case handling* [1]. This provides the insight that the approaches are not fully interchangeable, since they have specific motivations. In other words, drivers encompass broad and abstract goals (e.g., supporting communication and understanding), as well as very specifically desired and designed features (e.g., avoiding context tunneling).

2.2.2 Data integration

We identified four levels of data integration that data-centric process approaches pursue. The lowest three levels embrace the notion of a *data object* inspired by the object-oriented paradigm. With *object-oriented linking*, a process model is defined that captures the various activities and their ordering; each activity can invoke methods of an object to change the state of the object. This is the case in both *TriGSflow* [8] and *WASA₂* [25].

A tighter level of integration of data from the process is offered by those approaches that *identify explicit object life-*

cycles. In these approaches, similar to those on the previous level, the process interacts with objects by invoking methods/services to change the state of the object. However, each object also has a lifecycle, including its initial state, its final state, and possible transitions (through service calls to the object) that move the object from one state to another. *XFlows* [12], *object-centric process modeling* [20], and *object lifecycle compliant processes* [11] are the approaches in this category. Where the former uses the object lifecycle model to derive a classical, data-dependent activity-centric process model, the latter allows for the definition of both the object lifecycle model and the activity-centric process model, while ensuring that they comply with each other.

An even tighter way of integrating data is achieved by approaches that also offer a mechanism to synchronize different object lifecycles through *direct interactions* between objects. Note how this category assumes the capabilities of the previous one. Approaches with support for object interactions are *ACPM* [16], *DDPS* [15], *Object coordination nets* [27], *XDoc-WFMS* [9], *PBWD* [21], and *Philharmonic Flows* [10].

The strongest role for data is offered by those approaches where there is actually no explicit process anymore. The process steps follow from the availability of data objects, as production rules indicate how to produce new objects from existing objects or which operations are available based on the current data. The approaches that reside on this level are *case handling* [1], *PBWS* [24], and *document-driven workflows* [26].

Note that the levels in this dimension do not suggest any sense of superiority. Rather, the designers of the approaches emphasize the role of data to various degrees to offer support

for the design and execution of business processes. Since it is an open question to what extent the various approaches are successful in delivering on their promises, we will turn to that issue in the next section.

3 Evaluation method

On the basis of the presented literature review, we obtained a fairly good insight into the spectrum of data-centric process approaches. The main question that interests us in this paper, however, is how good the various approaches actually are. In other words, we are interested in their quality. The IEEE Standard Glossary of Software Engineering Terminology [19] proposes two alternative definitions of quality: (1) the degree to which a system, component, or process meets specified requirements; (2) the degree to which a system, component, or process meets needs or expectations of a customer or user. Both interpretations emphasize the relative nature of quality: There are requirements or expectations that at some point need to be confronted with reality. The second definition especially points out the role of the user in assessing quality.

3.1 Practitioner assessment

To make that quality notion operational, we aimed for the involvement of BPM practitioners. The basic idea was to let them try out the various approaches in a setting as realistic as possible and compare their experiences with what the approaches purported to bring. For practical reasons, we decided to focus here on DDPS, PBWD, and ACPM. This choice is to some extent arbitrary, but we did take two aspects into consideration. First of all, all three share the driver to increase the flexibility of business processes designed in this way. As may be recalled by the reader, this is the most popular driver behind data-centric process approaches (see Fig. 1). Second, all three approaches have been developed in an environment in which industry and practice interact (see Table 1). These aspects ensure that there is a common basis for comparison and that these approaches have to some extent already been tested out in practice, respectively. Furthermore, because they have been developed in such an environment, one would expect to find a high degree of fit between the quality claims as defined by the developers and how these are evaluated by users.

Apart from the similarities between the three selected approaches with respect to their drivers, they also display differences that make a more detailed comparison interesting. Their main difference lies in the central model that is used in each approach. DDPS puts the lifecycles (i.e., behavioral models) of process data central, while ACPM puts an object model of the artifacts at center stage. By contrast, PBWD builds on a data type model of these artifacts. An

illustration of the different approaches can be found in the examples given in the tutorials that were used in our evaluation and that we already referred to (see <http://is.ieis.tue.nl/staff/ivanderfeesten/Research/Data-centric/>). We also briefly illustrate the three approaches in the following subsection.

3.2 Three approaches

Data-Driven Process Structures (DDPS) use a standard *Data Model* to describe the object types involved in the process; each object type gets an *Object Lifecycle* (OLC) defining the states an individual object can progress through. A *Life-cycle Coordination Model* (LCM) combines the associations between object types defined in the static data model with the dynamic OLC models. In particular, the LCM specifies how events from one OLC relate to events from another OLC, for objects that are linked via an association from the data model. The data model, all OLCs, and the LCM are assumed to be reusable in different process contexts and therefore defined by technically skilled domain experts. A special characteristic of the method is that it provides a visual language not only for the data model and OLCs but also for the LCMs.

Our online tutorial supplement presents in detail an example from Müller et al. [14]. The example relates to a car navigation system with various subsystems. The data model involves just two classes (System and SubSystem) and one association (hasSubsys). Both System and SubSystem have a specific OLC. The LCM defines two external state transitions for association hasSubsys (one triggering a System state update based on a SubSystem state change and one vice versa). For the example, the data model is instantiated for a navigation system involving three subsystems (a main unit, a sound system, and a TV tuner). The structure and behavior of these component instances follow from the aforementioned static and dynamic models.

Product-Based Workflow Design (PBWD) describes all *data elements* involved in the process in a diagram similar to a *Bill of Materials*. *Leaf data elements* are initially given; *production rules* define which information of which elements has to be combined to obtain other elements; production rules can have additional attributes and conditions; *top elements* are the final data elements that shall be obtained.

Our online tutorial supplement presents in detail an example from [21]. The example relates to determining someone's suitability to become a helicopter pilot. Decisions are based upon psychological fitness, physical fitness, latest results of a suitability test from the last 2 years, and the quality of reflexes and eyesight. These decision aspects are modeled via primitive data elements. Associations are not modeled in the PBWD approach. The visual data model does include edges between data elements. These edges represent the production rules. In the given example, the top data element

is the final decision. Intermediate decision steps are represented as nodes that have a cost and time attribute. For the leaf elements in the example (elements that are acquired from external environment), no attributes of the production rules are specified; one can consider the cost and time for creating these elements to be zero. The example includes constraints such as “eyesight = bad.” Using the example model and its constraints, the best solution with respect to a performance indicator (cost or time) can be selected automatically.

In *Artifact-Centric Process Models (ACPM)*, a *business artifact* is a concrete, self-describing, and unique entity that corresponds to a key data object of the process. Artifacts have a unique identity (ID) that cannot be changed; the other attributes of artifacts can be changed at will. Each artifact has a *lifecycle model* which describes the states of an artifact from creation to disposal. The *information model* of the process describes all business artifacts and their relationships.

Besides artifacts, *services* are another key ingredient of ACPM. They make changes to one or more artifacts. Service changes are transactional. The four key aspects of a service are the Inputs, Outputs, Preconditions, and Effects (IOPE). Finally, *association rules* form the glue between services and events from different lifecycle models.

Our online tutorial supplement presents in detail an example based on [3]. The example relates to the provisioning, installation, maintenance, and general support for IT across multiple enterprise locations. The example covers seven artifact types (including Customer, Site, Schedule, Task, and Vendor). These artifacts are related via associations such as *BelongsTo* and *Supplies*. The example includes a service *Create_Schedule* which has three input artifacts and two output artifacts. The effect of this service includes setting status of the newly created schedule. The example also covers association rules in Event–Condition–Action (ECA) form.

It should be noted that the process specification on the artifact-centric approach has changed over the years, with the notable inclusion of Guard–Stage–Milestone lifecycles in 2011. We refer in our work to its original formulation in [16] and extension in [3].

In the following, we will describe how we established the claims for the three approaches on the one hand and how we confronted these with the experiences of real users on the other.

3.3 Claims

To establish the claims on the various approaches, we turned to the two highly cited papers that we also used for the classification in Sect. 2.2. To classify the various claims made in these papers and in the absence of a specific standard for process modeling approaches, we used the ISO/IEC 9126

software quality standard. This is a general-purpose standard for intangible assets, in particular software, which covers perspectives on both usefulness and ease of use. It includes the following categories:

1. *Functionality*: a defined objective or characteristic action of the approach.
2. *Usability*: the ease with which a user can learn to operate, prepare inputs for, and interpret outputs of the approach.
3. *Efficiency*: the degree to which the approach performs its designated functions with a minimum consumption of resources.
4. *Maintainability*: the ease with which the approach can be used to correct faults, improve performance or other attributes, or adapt to a changed environment.
5. *Portability*: the ease with which the approach can be modified for use in applications or environments other than those for which it was specifically designed.

Note that we left out *reliability*, since this quality aspect assumes sustained usage—a condition at odds with the workshop-like evaluation we selected.

Next, both papers for each approach were analyzed using a qualitative method borrowed from the toolkit of grounded theory [23]. The exact procedure follows the phases as suggested by Flick [6]. A first pass entails *open coding*, essentially consisting of tagging essential quotes on the quality of the approach with codes. In a second pass, similar or identical codes are grouped into concepts; these collections of codes describe the underlying idea on the basis of multiple codes. The final level in the coding hierarchy built up in this way is obtained when similar concepts are grouped; we set out to match these with the earlier defined categories. The used hierarchy of coding is in Table 2.

As can be seen, different approaches make different claims with respect to the same quality attributes, e.g., consider in the *Functionality* category how the underlying goal (driver) of the approach is claimed to be accomplished. At the same time, there are empty cells indicating that certain approaches do not make any claims with respect to that quality aspect, e.g., consider the understandability aspect in the *Usability* category for DDPS.

3.4 Workshops

The claims as identified in the preceding section were evaluated in a workshop setting, conducted at Eindhoven University of Technology in the start of 2013. Participation in these workshops was voluntary and solicited through invitations via the mailing list of the Dutch BPM Round Table.¹ In total, 29 people participated: 27 male versus 2 females,

¹ See: <http://is.ieis.tue.nl/research/bpmroundtable/>.

Table 2 Quality claims on the data-centric process approaches

| | Data-Driven Process Structures (DDPS) | Product-Based Workflow Design (PBWD) | Artifact-Centric Process Modelling (ACPM) |
|--|--|---|--|
| FUNCTIONALITY: Goal of method | [+] D1 Data-driven modelling of large process structures; ensuring correct coordination, reducing modelling efforts, and providing mechanisms for maintenance | [+] P1 Method that (re)designs a process reasoning from the desired outcome, without directly discussing how to achieve it | [±] A1 Representation usable by business people to analyse, manage, and control business operations [+] A2 Substantial new insights can be acquired by managers |
| How to achieve | [+] D2 Separation of data and process logic, leading to an enactable process structure | [+] P2 Analytical, clean sheet approach: rational and quantitative way of deriving an optimised process design | [+] A3 Declarative approach, that incorporates formality required for rigorous design and analysis |
| IT Support | [o] D3 IT support for automated creation and soundness checks | [o] P3 PDM used to steer a workflow execution: - basis for a process model | |
| USABILITY: Understandability | - | [±] P4 Created models are accepted by end users as valid and workable [+] P5 Based on existing concept (BOM) | [−] A4 Intuitive appeal to business managers, somewhat foreign to business process professionals [+] A5 Based on existing methods and techniques |
| User: ease of use | [−] D4 Intuitive integration of data and (sub-)processes enables users to instantiate and adapt model without process knowledge (by adapting data structure) | - | - |
| Expert: ease of use | [+] D5 Process experts require profound domain knowledge to create data model and LCM | [o] P6 Efforts required to collect data for PDM differs for every company [+] P7 Constructing PDM is manual task | [o] A6 Identifying artifacts is an iterative process that requires understanding of whole business process |
| EFFICIENCY: Different processes from same models | [+] D6 Instantiating different data structures and generating the respective data-driven process structures | [+] P8 Using different optimisation criteria, different process models can be created | [±] A7 Enabling specification of generic schema with multiple specialisations |
| Automated creation | [+] D7 Automated creation of process structure | [o] P9 Algorithms and PBWS can be used to - automatically generate process models or recommendations for carrying out operations | |
| Reusability | [+] D8 Standardising processing of objects in order to increase the reuse of process models and reduce modelling efforts | - | - |
| Reduce modelling efforts | [+] D9 Separation of data and process logic results in a reduction of modelling efforts | - | [o] A8 Separation of data management concerns from process flow concerns |
| Optimisation | - | [+] P10 Optimisation of the process by including - optimisation criteria | |
| MAINTAINABILITY: Changes | [+] D10 Changes can be made at data-level and are automatically translated into corresponding adaptations of the process | - | [−] A9 Changes can be made to conceptual flow and business workflow, while preserving the same Business Operations Model |
| FLEXIBILITY: Other environments | [+] D11 Usable in other environments, not specifically the engineering domain | [+] P11 Restricted to fields where clear concept of the products to be delivered exists | [+] A10 Usable in other environments, particularly in the use of consumed and non-consumed goods |

Legend: [+] = supported by participants, [±] = no consensus, [−] = participants contradict the claim, [o] = not addressed by any

20 professionals versus 9 students. Of the participating professionals, 15 assessed themselves to have intermediate expertise in process modeling and design, 5 were experts. The professionals were mostly consultants active in the BPM domain (13); the others had the following roles: software engineer (3), business analyst (2), process analyst (1), and process architect (1). Their average experience in the BPM field was a little less than 8 years. The students were all graduate students, with no practical BPM experience. They did

follow courses in their Industrial Engineering program on Modeling, Simulation, and Business Process Management on a graduate level at Eindhoven University of Technology. We consider these students good proxies for early-career BPM analysts. None of the participants had any previous experience with the selected approaches. While maintaining the same ratio of professionals versus students for each group, the participants were otherwise randomly distributed to three groups. Each group was assigned one single data-

centric process approach to evaluate. For each group, a single workshop was conducted, which lasted between 2 and 3 h. The phases that were gone through for each workshop were the same and will be described next.

The first phase of the evaluation encompassed (1) a tutorial to provide the participants with sufficient knowledge to evaluate a specific method and (2) exercises to let the participants apply their knowledge acquired in this way. Both the tutorials and the exercises for each method were created so that they would be self-explaining; no involvement of the research team during the workshops was planned or required. Prior to the workshop, a first draft of the materials was created based on the existing literature. These materials were then reviewed by the proposers of the method (see Table 1), who graciously provided their support to ensure that the contents of the tutorial were clear and correct. After adjustments were made and additional information that was received from the developers was included, two graduate students not participating in the evaluation pretested the materials leading to minor textual adjustments.

The second phase of the evaluation consisted of having the participants fill out a questionnaire on the usability of the approach they had studied and applied. We used the questions that were developed as part of Moody's Method Evaluation Model [13]. This is a broad-brush evaluation method that translates the determinants from the well-known technology acceptance model [5] to IS design methods. The main purpose of this phase was to establish whether professionals and students differed significantly with respect to their perceptions of the methods. After all, this would call for a different valuation of their input.

The third and final phase of the workshop encompassed a semi-structured group discussion [18]. The following questions were used to guide the discussions:

1. Were there unclear aspects regarding the method or its application?
2. What do you think are strong points of this approach?
3. What is in your opinion a limitation or weakness of this approach?
4. Was the method complete? If not, what was lacking?
5. How versatile do you think the method is?
6. Do you have any suggestions for improving the method?

The different categories of the ISO/IEC 9126 standard we discussed in Sect. 3.3 were used to both categorize the answers and to solicit follow-up discussions to cover each of these.

The discussions were all recorded. When statements were made during the workshop on quality aspects of the methods, the moderator followed up to determine the support for such statements among the participants. The workshops were concluded with a brief explanation of the solutions to the exercises.

3.5 Matching

To match the claims with the workshop results, the following procedure was followed. The recorded discussions were all split up into excerpts. These are quotes made by the participants, capturing their statements and responses to each other. Each quote was identified by a number, which was used to locate the specific quote. After the quotes were captured, they were coded in a similar fashion as explained in Sect. 3.3 and then categorized using the quality attributes explained in the same section.

After matching the quotes from the discussions to the claims, it became possible to determine to what extent the various claims were supported. We distinguished four options for each claim, with a corresponding color code (see Table 2): green +: The claim is supported by the participants; orange ±: There is no consensus among the participants; red -: Participants contradict the claim; gray o: The claim is not addressed.

A claim was considered to be contradicted if less than 35% of the participants in a workshop supported the claim; if support was between 35 and 65%, it was decided that no consensus was reached; finally, a claim was considered to be supported if at least 65% of the participants supported the claim. In addition, it occurred that one statement supported and another one contradicted a claim; in such a situation, we decided that no consensus was reached. In the next section, we will reflect on the results obtained this way.

4 Results

The first result that is presented here are the outcomes of Moody's Method Evaluation Model. Recall that this method helps us to see whether the students and professionals have fundamentally different views on the quality of the evaluated methods. It also gives us a first idea of the participants' perception of the quality of data-centric methods in general and their usability in particular. A standard analysis was applied to test the internal consistency of the used evaluation questions, showing that both tested variables *perceived ease of use* and *perceived usefulness* could be kept with values exceeding the generally accepted threshold of 0.7 [17]. A graphical representation of the evaluations in the form of boxplots is provided in Fig. 2. A boxplot displays the median of a distribution as a horizontal dash inside a box. The 25% of the data points right above this median are in the top part of this box; the 25% of data just below the median are in the lower part of the box. The whiskers show the greatest and lowest values when excluding outliers—the latter are presented individually.

The figure provides the following insights. First of all, the distributions of the perceptions—scored on a scale between

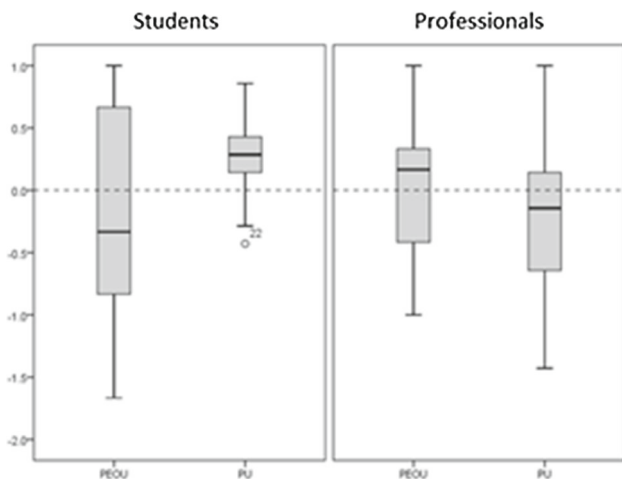


Fig. 2 Perceived ease of use (PEOU) and perceived usefulness (PU)

−2 (very negative) and +2 (very positive)—center around the neutral stance, which is indicated with a dashed line. Second, the perceptions of professionals and students do not differ widely. While professionals seem to be slightly more positive with respect to the ease of use of the data-centric approaches, students are more so with respect to their usefulness. An additional one-way ANOVA test indicated that there are no significant differences between the perceptions of professionals and students. Because of this, the remaining analysis is provided for each method separately, aggregated over all participants per method (students and professionals). We will refer to the various claims as listed in Table 2.

4.1 Data-Driven Process Structures (DDPS)

4.1.1 Functionality

The participants agreed that the approach reduces modeling efforts and provides mechanisms for maintenance (D1). The reduction in modeling efforts is realized by model reuse and the automated generation of the workflow, while maintainability is realized by separating data and process logic (D2). No automated tools were incorporated in the workshop, which leaves D3 neither supported nor contradicted.

4.1.2 Usability

Although participants do see the benefits when users only have to instantiate a data model, they indicated that the models were likely to be too complex for managers and other users. This contradicts claim D4, which suggested an intuitive integration of data and processes. Specifically, they question the interpretability of the abstract models for users.

Claim D5 was supported: Profound domain knowledge will be required to create the data model and the LCM.

4.1.3 Efficiency

All efficiency claims found support among the participants. Due to the separation of data and process logic, a data model can be used for creating different data structures and their corresponding DDPS (D6 and D9). The automated creation of a process structure based on the data model and the LCM makes the method efficient in both time and other resources (D7). Finally, the participants identified the reusability of lifecycles and data models as one of the major contributors to the method's efficiency, which supports D8.

4.1.4 Maintainability

The participants acknowledged that adjustments to a model or process can be made in a single place and can be automatically translated to a new process structure (D10).

4.1.5 Portability

It was easy to imagine for the participants that DDPS would be applicable in other domains, which supports D11. Specifically, the administrative domain was mentioned as a candidate.

4.1.6 Additional findings

In the discussion, the participants explicitly mentioned the use of hierarchy as a positive usability element in the approach. However, recurring in the discussion also was that large process structures were difficult to read and interpret.

4.2 Product-Based Workflow Design (PBWD)

4.2.1 Functionality

The PDM used in PBWD is, conform to P1, perceived as a good basis for redesign: It is orderless, points out process dependencies, and provides insights in important aspects of the (re)designed process. In addition, since the method is orderless and uses very explicitly defined rules, an objective (re)design can be created (P2). P3, which considered the runtime concern of workflow execution, was not addressed in the discussion.

4.2.2 Usability

According to some participants (5 in total), a PDM is usable for customers. Also, especially in systems that include many

rules and calculations, a PDM would be highly understandable. The remainder of the participants (4 in total), however, perceived the PDM as too complex for end users. Therefore, there is no consensus for P4, which argues that created models are accepted by end users as valid and workable. Support for P5 was established, as the similarities to a Bill of Materials were indeed identified (P5). Efforts to collect the data required for a PDM were neither acknowledged nor refuted (P6); the discussion focused on difficulties in constructing a correct and complete PDM. The participants did indicate that the creation of a PDM is a task of a business analyst, able to derive and use abstract models. This indicates that they do see it as a manual task, cf. P7.

4.2.3 Efficiency

The participants mentioned various criteria, e.g., cost and time, which can be used to optimize different processes on the basis of a PDM. This supports P8, which suggested this multi-criteria aspect. Since no tool support was involved in the workshop, it was not possible to find support or counter-support for P9. P10 was deeply discussed and supported and strongly acknowledged: It was clear for the participants that process optimization is possible using the approach.

4.2.4 Maintainability

No claim was found regarding maintainability for this approach.

4.2.5 Portability

The participants of the workshop perceived the method as broadly applicable. They did not specifically acknowledge the restriction of a process that includes a clear concept of the product as the developers did, although it was mentioned that real problems would likely not to be too straightforward to tackle. Combining these two viewpoints, P11 was determined as being supported.

4.2.6 Additional findings

The discussion on the usability of the PDM, which was already perceived as being difficult by the participants, focused on a particular element sometimes found in a PDM: alternative routings. This construct, especially in large models, adds to a lower comprehension of such a model. Also, using a PDM requires an alternative way of thinking (compared to activity-centric methods), which is also perceived as difficult by some of the experts and as potentially constructive to reason about optimizing the process by others.

4.3 Artifact-Centric Process Models (ACPM)

4.3.1 Functionality

Claim A1 did not find full support: Although most participants stated that the combination of artifacts and lifecycles helped to understand the interaction of different artifacts in a system, most of them did *not* see this as implying it could be used to analyze, manage, and control business operations. Participants did acknowledge the method for being able to provide new insights in the processes (A2). Furthermore, using the structured approach of this method, knowledge, roles, and rules, required for rigorous design and analysis, can be captured, which directly supports A3.

4.3.2 Usability

The supposedly intuitive appeal of the method (A4) was contradicted by the participants: The artifacts were seen as non-intuitive (for both business users and experts), and reasoning using the proposed method was perceived to be rather confusing. The use of existing and known methods was seen as an advantage, which supported claim A5. Claim A6, which is related to the understanding of the whole process, was not addressed: None of the participants had thorough domain knowledge on the process used in the exercises.

4.3.3 Efficiency

Most participants acknowledged that a process could reuse already defined artifacts for a new purpose, although less directly as formulated in the claim. This provides partial support for A7. The separation of data management concerns from process flow concerns (A8) remained unaddressed in the workshop.

4.3.4 Maintainability

In the eyes of the participants, maintaining the models would be fairly difficult, since adaptations in artifacts or their lifecycles often result in making changes in multiple places. We found this to refute claim A9, which suggested that some models would be stable.

4.3.5 Portability

The method was perceived by all participants as very broadly applicable, i.e., full support for A10. While the developers of the method further specified this to consumable and non-consumable goods, the participants did not further elaborate on this statement.

4.3.6 Additional findings

In addition to the interaction of artifacts and lifecycles, which are related to claim A1, the possibility to incorporate hierarchy in the method was mentioned as an additional functional strength. With respect to usability, the consensus was that the required models were difficult to determine and of unclear value.

5 Discussion

Given the evaluation of the claims that were presented in the previous section, a number of insights stand out. First, for all approaches, the claims with respect to their functionality find full support in most cases (D1, D2, P1, P2, A2, and A3) and partial support for one claim (A1).² A similar pattern appears with respect to efficiency: Most claims find full support (D6, D7, D8, D9, P8, and P10) and one claim finds at least partial support (A7). These positive evaluations of functionality and efficiency claims seem to suggest that the workshop participants agree with the proposers of the methods that the necessary models and the resulting process support can be efficiently generated for their specific purpose. In that sense, the quality of the data-centric process approaches is clearly acknowledged. Their potential is also reinforced by the positive evaluations of the claims with respect to the portability of all three approaches to other domains (D11, P11, and A10).

Second, across all approaches, the usability claims come off badly. For DDPS and ACPM, participants refute claims that the integration between process and data in the models is so intuitive that business users can easily make sense of these (D4 and A4). PBWD, too, encounters only partial support for the claim that end users accept the created models as valid and workable (P4). This points at a disconnect between what the proposers of the methods claim and what is actually experienced by users. What may be behind this is that the proposers of methods developed a positive bias toward their own techniques, either through their prolonged exposure to these techniques or vested interests. Moreover, it may also offer, albeit contentiously, an explanation for the limited uptake of these data-centric process approaches: The various modeling techniques and models are not that easy to use, at least not without IT support. This could be interpreted as a call to action for developers of data-centric approaches to consider the human factor more than is currently the case. We believe that the setup of the workshop-centered evaluation approach as used in this paper could serve as a source for inspiration to generate explicit feedback on usability aspects.

² Note that this excludes claims with respect to IT support (D3, P3), since these were not incorporated in the workshops.

Third, the workshop discussions helped identify improvement opportunities for all three approaches. For DDPS, the links between hierarchical levels and objects in the LCMs need to be better supported visually. A separate view on these relations may be helpful. For PBWD, it seems proper to rethink how alternative ways to create the value of a single data element are visualized in a PDM: It was universally recognized as a confusing aspect. One direction could be to explicitly label each of the incoming arcs as alternatives. For ACPM, the main point of criticism is related to the maintainability of the conceptual flow and business workflow implementation (A9). The challenge here is to find better ways to propagate adaptations in artifacts or their lifecycles toward other places, instead of relying on manual changes. Tool support seems crucial for this task.

6 Conclusion

We departed from the observation that the emphasis in the BPM community seems to be on developing approaches, without much attention for the needs and preferences of the individual who is modeling, e.g., the usability of such approaches. Our study attempts to shed some light on this, in the specific context of data-centric process approaches.

Our work has led to the following contributions. First and foremost, we found that the usability aspects of all considered approaches are a source of concern. This has led to the insight that the usability aspects of process design approaches are crucial and, in the perception of groups representative of actual users, leave much to be desired. In fact, we show that usability—the modeler's interest—takes the backseat. In that sense, our research can be seen as a wake-up call for process modeling notation designers to take the usability side, and as such, the interest of the human modeler, more seriously. It is not sufficient that new process design approaches are continuously launched by the community without serious consideration for the intended users of such methods.

Second, our research shows how to set up an evaluation of a new modeling/design approach, which can be considered an additional contribution. The purpose of this evaluation approach is to determine whether claims behind approaches could be substantiated in a setting where they were actually used by independent users. Specifically, three data-centric approaches have been evaluated by groups of professionals and students. These approaches are similar in that they share drivers, including the most common one (increased flexibility), and reside on the same level of data–process integration. We found improvement opportunities for each of the three approaches.

Third, before the above evaluation, a classification of the available approaches has been made, which uses the

driver behind the approach as well as the level of data integration in the process approach. This classification is considered valuable to understand the exact positioning of each approach and to determine how prominent they treat the data perspective.

Our findings must be seen against a number of limitations of the underlying research approach. First, our view on the literature is by definition limited in time and scope. We do not claim completeness of the presented classification, but settle for a high level of comprehensiveness. Also, we acknowledge that many of the approaches under consideration are still in development, which makes our efforts somewhat like “shooting at a moving target.” Second, the workshops have been constrained with respect to the number of participants, the training that could be provided to the participants, the number of cases that were dealt with and their realism, and the time available to work on these cases. As discussed in the paper, we have taken steps to mitigate negative effects. Notably, the proposers themselves validated the tutorials and exercises. In addition, we tested for undesirable effects of mixing experts and novices in the evaluation. Despite the potential biases of empirical research in general and those of this specific investigation in particular, we believe that it is important to better understand the strengths and weaknesses of data-centric process approaches as experienced by actual users of them. If anything, we hope that our work contributes to the notion in the BPM field that claims call for validation.

The presented research can be strengthened and extended in various ways. A straightforward extension is toward the inclusion of more process approaches. A more sophisticated point worth considering is that in the current setup it was not possible to determine the quality of the various approaches *in comparison with each other*. Note that different groups of participants evaluated different approaches and that the materials they were provided with on the cases were different. It would be of much interest to develop and execute a setup in which the approaches would be more directly compared. This is challenging in its own right since the various approaches require partly different inputs. In addition, it would require the intense involvement of the same participants over a much prolonged period of time. A step beyond these possible extensions of our study would be to compare data-centric process approaches with more conventional ones. Related to this, it would be interesting to further study the relation between data-centric approaches and usability, e.g., whether approaches following this “school” are by definition less intuitive to users. Finally, we believe that the tool aspect in our investigation has been underexposed. In our plans to follow up on this research, we are looking for ways to have participants use the specific tool sets to develop and generate the models under consideration.

Acknowledgements The authors thank all respondents involved in the evaluation study and acknowledge the valuable input by the proposers of the respective approaches.

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

1. Van der Aalst, W.M.P., Weske, M., Grünbauer, D.: Case handling: a new paradigm for business process support. *Data Knowl. Eng.* **53**(2), 129–162 (2005)
2. Bandara, W.: Process modelling success factors and measures. Ph.D. thesis, Queensland University of Technology (2007)
3. Bhattacharya, K., Hull, R., Su, J.: A Data-Centric Design Methodology for Business Processes, pp. 503–531 (2009). <http://www.igi-global.com/bookstore/Chapter.aspx?TitleId=19707>
4. Cooper, H.M.: *Synthesizing Research: A Guide for Literature Reviews*, vol. 2. Sage, California (1998)
5. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: a comparison of two theoretical models. *Manage. Sci.* **35**(8), 982–1003 (1989)
6. Flick, U.: *An Introduction to Qualitative Research*. Sage, California (2009)
7. Henriques, R., Silva, A.: Object-centered process modeling: principles to model data-intensive systems. In: Muehlen, M., Su, J. (eds.) *Proceedings of the Business Process Management Workshops. Lecture Notes in Business Information Processing*, 66th edn, p. 683. Springer, Berlin (2011)
8. Kappel, G., Rausch-Schott, S., Retschitzegger, W.: A framework for workflow management systems based on objects, rules and roles. *ACM Comput. Surv.* **32**(1es), 27 (2000)
9. Krishnan, R., Munaga, L., Karlapalem, K.: XDoc-Wfms: a framework for document-centric workflow management system. *LNCS* **2465**, 0348–362 (2002)
10. Künzle, V., Reichert, M.: PHILharmonicFlows: towards a framework for object-aware process management. *J. Softw. Maint. Evol. Res. Pract.* **23**(4), 205–244 (2011)
11. Küster, J.M., Ryndina, K., Gall, H.: Generation of business process models for object life cycle compliance. In: *BPM 2007. LNCS*, vol. 4714, pp. 165–181 (2007)
12. Marchetti, A., Tesconi, M., Minutoli, S.: XFlow: An XML-Based Document-Centric Workflow. In: *WISE. LNCS*, vol. 3806, pp. 290–303. Springer (2005)
13. Moody, D.L.: The method evaluation model: a theoretical model for validating information systems design methods. In: *Proceedings of the 11th European Conference on Information Systems*, pp. 16–21. Naples, Italy, June (2003)
14. Müller, D., Reichert, M., Herbst, J.: A new paradigm for the enactment and dynamic adaptation of data-driven process structures. In: *Proceedings 20th International Conference on Advanced Information Systems Engineering (CAiSE'08)*, Montpellier, France. *Lecture Notes in Computer Science*, vol. 5074, pp. 48–63. Springer Verlag, London (June 2008)
15. Müller, D., Reichert, M., Herbst, J.: Data-driven modeling and coordination of large process structures. In: *CoopIS 2007. LNCS*, vol. 4803, pp. 131–149. Springer (2007)
16. Nigam, A., Caswell, N.S.: Business artifacts: an approach to operational specification. *IBM Syst. J.* **42**(3), 428–445 (2003)
17. Nunnally, J.C.: *Psychometric Theory*. McGraw-Hill, New York (1978). (1978)

18. Qu, S.Q., Dumay, J.: The qualitative research interview. *Qualitative Res. Acc. Manage.* **8**(3), 238–264 (2011)
19. Radatz, J., Geraci, A., Katki, F.: IEEE standard glossary of software engineering terminology. *IEEE Std 610121990, 121990* (1990)
20. Redding, G., Dumas, M., ter Hofstede, A.H.M., Iordachescu, A.: A flexible, object-centric approach for business process modelling. *Serv. Oriented Comput. Appl.* **4**(3), 191–201 (2010)
21. Reijers, H.A., Limam, S., van der Aalst, W.M.P.: Product-based workflow design. *J. Manage. Inf. Syst.* **20**(1), 229–262 (2003)
22. Rodriguez-Priego, E., Garcia-Izquierdo, F., Rubio, A.: Modeling Issues: a Survival Guide for a Non-expert Modeler. In: Petriu, D., Rouquette, N., Haugen, Ø. (eds.) *Model Driven Engineering Languages and Systems, Lecture Notes in Computer Science*, vol. 6395, chap. 26, pp. 361–375. Springer, Berlin, Heidelberg (2010)
23. Strauss, A., Corbin, J.: Grounded theory methodology. *Handbook of qualitative research* pp. 273–285 (1994)
24. Vanderfeesten, I.T.P., Reijers, H.A., van der Aalst, W.M.P.: Product-based workflow support. *Inf. Syst.* **36**(2), 517–535 (2011)
25. Vossen, G., Weske, M.: The WASA2 object-oriented workflow management system. In: *SIGMOD Conference 1999*. pp. 587–589. ACM Press (1999)
26. Wang, J., Kumar, A.: A framework for document-driven workflow systems. In: *BPM 2005. LNCS*, vol. 3649, pp. 285–301. Springer (2005)
27. Wirtz, G., Weske, M., Giese, H.: The OCoN approach to workflow modeling in object-oriented systems. *Inf. Syst. Front.* **3**(3), 357–376 (2001)



Hajo A. Reijers is a full professor in Business Informatics at Vrije Universiteit Amsterdam and a part-time full professor in Information Systems at Eindhoven University of Technology. He worked as a management consultant for Deloitte and led an R&D group within Lexmark. He published over 150 scientific papers, chapters in edited books, and articles in professional journals on topics such as business process management, workflow technology, process mining, conceptual modeling, and simulation.

conceptual modeling, and simulation.



Irene Vanderfeesten is an assistant professor in Business Process Management and Information Systems at Eindhoven University of Technology (TU/e). She received her M.Sc. degree in Computer Science and her Ph.D. degree in Industrial Engineering from TU/e in 2004 and 2009, respectively. Before joining the Information Systems group of the department of Industrial Engineering and Innovation Sciences at TU/e in 2010, she also worked as an IT consultant

in the banking and insurance sector. Her current research interests include business modeling; business process redesign and improvement; workflow management; and human aspects of business process management and information systems.



Marijn G. A. Plomp is an assistant professor at the Knowledge, Information and Innovation (KIN) group at Vrije Universiteit Amsterdam. His research focuses on the combination of technology and organization. He is also Programme Manager at the Amsterdam Center for Business Analytics (ACBA). Marijn holds a Ph.D. in Information Systems from Utrecht University, based on his thesis “Maturing Interorganisational Information Systems.” Before joining

VU University, Marijn worked as researcher/consultant. His research appears in various international conferences and journals, such as *Electronic Markets*, *International Journal of Information Technology and Management*, and *Supply Chain Management: An International Journal*.



Pieter Van Gorp is conducting research on personal health data as an economic asset. He considers health records as well as evidence-based workflow and decision support models as examples of this asset with underused economic potential. Regarding health records, he has performed research on MyPHRMachines, a platform for securely analyzing personal health records in the cloud. Regarding workflow models, he has contributed novel transformations for UML, BPMN and Petri-Net models. Van Gorp is also a facilitator of reproducible research (e.g., via the SHARE20.eu cloud). Since 2008, he is an assistant professor in the School of Industrial Engineering at Eindhoven University of Technology. Previously he held a postdoc position at the University of Antwerp, where he also obtained his Ph.D. degree in Software Engineering.

Since 2008, he is an assistant professor in the School of Industrial Engineering at Eindhoven University of Technology. Previously he held a postdoc position at the University of Antwerp, where he also obtained his Ph.D. degree in Software Engineering.



Dirk Fahland is assistant professor at the Eindhoven University of Technology researching in the area of distributed systems, and he received his Ph.D. in Computer Science from the Humboldt-Universität zu Berlin, Germany, and the Eindhoven University of Technology, the Netherlands, in 2010. His research interests include distributed processes and systems built from distributed components for which he investigates modeling systems (using process

modeling languages, Petri nets, or scenario-based techniques), analyzing systems for errors or misperformances (through verification or simulation), and process mining/specification mining techniques for discovering system models from event logs. He particularly focuses on distributed system with multi-instance characteristics and their syn-

chronizing and interacting behaviors. His results appeared in journals such as *Software & Systems Modeling*, *The Computer Journal*, *Data and Knowledge Engineering*, and *Information Systems*.



Wim L. M. van der Crommert is a business analyst at Precedence B.V., a process oriented consultancy company. He worked at the design of large projects in different multinationals in Europe and facilitated in the implementation of these projects. He holds an M.Sc. in Operations Management and Logistics from Eindhoven University of Technology. His main interests are business process redesign, workflow management systems, and change management.



H. Daniel Diaz Garcia is an Engagement Manager of Operations and Business Support Systems (OBS) at Ericsson Mexico. He received his B.Sc. degree in Telematics Engineering from Mexico Autonomous Institute of Technology (ITAM) and M.Sc. degree in Business Information Systems from Eindhoven University of Technology (TU/e). Hector leads OBS solutions design for major telecommunications operators in Latin America and the Caribbean. His solution

domains include billing, charging, catalog-driven order management, customer relationship management systems, and business process design of telecom operations.